

147838

Schreiber, David

From: Ramirez, Delia
Sent: Thursday, March 10, 2005 4:09 PM
To: Schreiber, David
Subject: case 10/034,985

Hi,

I would like to request the following search : SEQ ID NO:1 and 2 in the protein databases (commercial and interference).

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

THIS PAGE BLANK (USPIC,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:36:45 ; Search time 183.5 Seconds
(without alignments)
4756.237 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcacatcacatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 2793840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US10034985/runat_10032005_164243_12058/app query.fasta_1.1479
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQRY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Result. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query
Result

No.	Score	Match	Length	DB	ID	Description
1	2302	98.0	440	9	US-09-777-566A-2	Sequence 2, Appli
2	2302	98.0	440	9	US-09-866-379-2	Sequence 2, Appli
3	2302	98.0	440	14	US-10-034-985-2	Sequence 2, Appli
4	2302	98.0	440	15	US-10-430-356-2	Sequence 2, Appli
5	2302	98.0	440	15	US-10-601-313-2	Sequence 2, Appli
6	2235	95.1	432	9	US-09-866-379-8	Sequence 8, Appli
7	2235	95.1	432	14	US-10-156-660-4	Sequence 4, Appli
8	2235	95.1	432	15	US-10-282-122A-43351	Sequence 43351, A
9	2235	95.1	432	15	US-10-601-319-8	Sequence 8, Appli
10	2235	95.1	432	16	US-10-472-317-41	Sequence 41, Appli
11	2221	94.6	432	15	US-10-284-962-3	Sequence 3, Appli
12	2221	94.6	432	15	US-10-284-962-14	Sequence 14, Appli
13	2221	94.6	432	14	US-10-266-041-1	Sequence 1, Appli
14	2197	93.5	432	15	US-10-284-962-5	Sequence 5, Appli
15	2188	93.1	436	14	US-10-156-660-2	Sequence 2, Appli
16	2171	92.4	432	15	US-10-601-319-10	Sequence 10, Appli
17	2165	92.2	432	14	US-10-021-723A-16	Sequence 16, Appli
18	2141	91.1	430	9	US-09-866-379-10	Sequence 10, Appli
19	2070	88.1	412	14	US-10-334-672-1	Sequence 1, Appli
20	2070	88.1	412	14	US-10-334-671-1	Sequence 1, Appli
21	2066	88.0	410	14	US-10-021-723A-13	Sequence 13, Appli
22	2066	88.0	410	14	US-10-021-723A-15	Sequence 15, Appli
23	960.5	40.9	441	14	US-10-021-723A-12	Sequence 12, Appli
24	943.5	40.2	441	15	US-10-282-122A-77792	Sequence 2, Appli
25	935.5	39.8	441	14	US-10-021-723A-2	Sequence 2, Appli
26	932.5	39.7	420	14	US-10-021-723A-4	Sequence 4, Appli
27	760	32.4	476	14	US-10-021-723A-10	Sequence 10, Appli
28	611.5	26.0	318	14	US-10-021-723A-14	Sequence 14, Appli
29	476.5	20.3	409	14	US-10-021-723A-8	Sequence 8, Appli
30	472.5	20.1	421	14	US-10-021-723A-6	Sequence 6, Appli
31	195	8.3	426	15	US-10-257-174-44	Sequence 4, Appli
32	195	8.3	426	15	US-10-343-357-6	Sequence 6, Appli
33	126	5.4	3074	15	US-10-282-122A-65531	Sequence 5, Appli
34	121.5	5.2	19723	15	US-10-084-846A-5	Sequence 5, Appli
35	121	5.2	465	16	US-10-662-914-10	Sequence 10, Appli
36	120.5	5.2	19695	15	US-10-084-846A-3	Sequence 3, Appli
37	119	5.1	439	14	US-10-062-848-3	Sequence 3, Appli
38	119	5.1	439	16	US-10-776-104-3	Sequence 3, Appli
39	119	5.1	440	15	US-10-442-538-106	Sequence 106, App
40	119	5.1	440	15	US-10-442-538-128	Sequence 128, App
41	119	5.1	449	15	US-10-442-538-151	Sequence 151, App
42	119	5.1	449	16	US-10-776-104-12	Sequence 12, Appli
43	119	5.1	449	16	US-10-062-848-12	Sequence 8, Appli
44	119	5.1	465	13	US-10-083-452-8	Sequence 78, Appli
45	119	5.1	465	14	US-10-062-848-78	Sequence 78, Appli

ALIGNMENTS

RESULT 1
US-09-777-566A-2
; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERGA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT FILING DATE: 2001-06-11
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-04-13
; PRIOR FILING DATE: 1999-03-01
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2
 ; LENGTH: 440
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1323)
 ; OTHER INFORMATION: n is any nucleotide
 US-09-777-566A-2

Alignment Scores:

Pred. No.: 5,44e-205 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-777-566A-2 (1-440)

```

QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCCGCAATCTGCA 60
DB 1 MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGTTGATGTGTCAGTCGTCATGGT 120
DB 21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerA-gHisGly 40
QY 121 GTGCTGCTCAACCAAGGCGCAGCAACTGATGAGGATGTACCCCGACGATGCGCA 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTppPro 60
QY 181 ACCTGGCGGTAAATAGCTGGCTGACACCGCGGTTGGTGTGCTGAATCGCTATCTC 240
DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
QY 241 GGACATTACCAACCCAGCGCTGTGTAGCCGACCGATTGTCGCGAAAAAGGGGTGCGCG 300
DB 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTTCGGATATTGCTGATGTGACGAGCGTACCCCGTAAACAGGCGAA 360
DB 101 GlnSerGlyGlnValAlaIleAlaLeuValAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCCCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATTCCTCTAAACTGGCGTTTGCCAACTGGATAACGCG 480
DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGGATCTCAGCAGGCGAGAGGGTCAATTGCTGACTTTACCGGCAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTCCGCACTGCAACGGGTGCTTAATTTCCGCAATCAACTGTCG 600
DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAACAGGACGAAAGCTGTTTCAATTAACGAGGATTAACCATCGGAATC 660
DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTAGCGCGCAATGCTCTAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTGCAACAGCACAGGATCCGAGGATCCGCGGCGGTGGGAAGATCACC 780
DB 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACACGAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
  
```

```

DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
QY 841 CGCACCCGAGAGTTCCCGCGCAGCCCGCCACCCCGTATTGATTTGATTCATGGCAGCG 900
DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACGCAAAAACAGGCGGTATGGTGACATTACCCACTTCAGTACTG 960
DB 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCAGCTGAGACTCACTGG 1020
DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGCTCAGCGGATACACGCGCAGGTTGGTGAACCTGGTTGTAACGCTGG 1080
DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCCATAACACGACGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACGCGCTGTCTAATAATACGCGCCCGGAGAGGTGAAACTGACC 1200
DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGGATGCAAGAGCGAAATGCGCAGGCGATGTGTCGTTGCGCAGGTTTTACGCA 1260
DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACGACGATACCGCGGTGAGTTTGAGATCTCATCACCATCACCATCAC 1320
DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
  
```

RESULT 2

US-09-866-379-2
 ; Sequence 2, Application US/09866379
 ; Patent No. US20020136754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KRETZ, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT APPLICATION NUMBER: US/09/866,379
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: US 09/580,515
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 440
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; US-09-866-379-2

Alignment Scores: 5,44e-205 Length: 440
 Pred. No.: 2302.00 Matches: 440
 Score: 100.00% Conservatives: 0
 Percent Similarity: 100.00%


```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-866-379-2 (1-440)

Qy 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetlysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Qy 61 TTCGCTCAGAGTAGCGGAGCTCAAGCTGGAAGTGTGGTATTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLeuGlnSerValIleValSerArgHisGly 40

Qy 121 GTGCGTCTCAACCAAGCCACCAACTGATGAGGATGTCAACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Qy 181 ACCTGGCGGTAAACTGGGTGGCTGACACCCGCGGTGTGAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80

Qy 241 GGACATTACCAACCCAGCGTCTGCTAGCCAGCGATTGCTGGCGAAAGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100

Qy 301 CAGTCTGCTCAGGTCGCGATTATTGCTGATCTGACGAGCGTACCCGTAAACACAGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120

Qy 361 GCCTTCGCGCGCGGTGGCACTGCTGTCATTAACCGTACATACCCAGCGCATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140

Qy 421 TCCAGTCCCGATCCGTTATTATCTTAATCTTAATAACTGCGTTCCCACTGGATAACGG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

Qy 481 AACGTGACTGACGGATCCTCAGCAGGCGAGGGTCAATTGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180

Qy 541 CGGCAACCGCGTTTCGCAACTCGAATCGGCTGCTTAATTTCCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

Qy 601 CTTAAACGTGAGAAACAGCAGCAAGCTGTTCATTACGAGCGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Qy 661 AAGGTGAGCGCGCAATGCTCTCAATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

Qy 721 GAGATATTCTCTCGCAACAGCAGCGAATGCGGAGCGCGGTGGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260

Qy 781 GATTACACAGTGGAGACCTTCTGATTAAGTTTGATTAACCGCGCAATTTTATTCTCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280

Qy 841 CGCAGCGCAGAGTTGCCGCGCGCCACCGCTTATGATTGATTGATCATCGCAGCG 900
Db 281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300

Qy 901 TTGACGCGCCATCCACCGCAAAAAACAGCGGTATGTTGATTAACCCACTTCACTAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

Qy 961 TTTATTCGCGACACGATCTAAATCTGCGCAATCTCGCGCGCGCTGAGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340

```

RESULT 3

```

US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-034-985-2

```

```

Alignment Scores:
Pred. No.: 5,44e-205 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-034-985-1 (1-1323) x US-10-034-985-2 (1-440)

Qy 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetlysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Qy 61 TTCGCTCAGAGTAGCGGAGCTCAAGCTGGAAGTGTGGTATTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLeuGlnSerValIleValSerArgHisGly 40

Qy 121 GTGCGTCTCAACCAAGCCACCAACTGATGAGGATGTCAACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Qy 181 ACCTGGCGGTAAACTGGGTGGCTGACACCCGCGGTGTGAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80

```

```
QY 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGATTGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTCCGGATTATTGCTGATGTCACGAGGTACCCGTAACAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGCTGGACCTGACTGTGCATTAACCTGCAATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATTCCTCTAAACCTGCGCTTTGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGGATCCTCAGAGGCGAGAGGGTCAATTGCTGACTTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTTCGCGAACTGGAACGGTGTCTTAATTTCCGCAATCAAACTTCTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCATTAAACGAGGCAATTACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTAGCGCGCAATGCTCTAATTAAACGGTCCGTTAAGCTTCGGATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCCGCAACAGCAGCAGGATGCGGAGCGCGGTGGGAAGATCAACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglyArgIleThr 260
QY 781 GATTACACACGAGTGGAAACACCTTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACCGCAGAGTTGCGGAGCGCGCGCCGACCCCGTTATTGGATTGATCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCATGCTGTGATTCACATPACCACCTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGACACGATACCTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGTCAAGCCGGATAACACCGCGCAGTGGTGAACCTGGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCGATACAGCCACTGGATTTCAGTTTCGCTGGTCTCCAGCTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGGTGATAAAGCCCGCTGTCTAATTAATACGCCGCCGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGATGTGAAGACGGAATTCGCGAGGGCATGTGCTGGTGGCAGGTTTACGCCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACACCATACCGCGTGCAGTTTGTAGATCTCATCCATCCATCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

```
US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-04-13
; PRIOR FILING DATE: 1997-08-13
; PRIOR FILING DATE: 1997-08-13
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2
```

Alignment Scores:

Pred. No.:	5,44e-205	Length:	440
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.00%	Indels:	0
DB:	15	Gaps:	0

US-10-034-985-1 (1-1323) x US-10-430-356-2 (1-440)

```
QY 1 ATGAAAGCGCATCTTAATCCCATTTTATCTCTGATTCCTGTTAATCCGTTAATCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTTCAGTCTGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGCTCTCAACCAAGGCCACGCACTGATCAGAGTGTACCCAGAGCATGCCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGGTAAACTGGGTTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTTACCAACGCGAGCTCTGGTAGCCGAGATTGCTGGCGGAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTCCGGATTATTGCTGATGTCGAGCGGTACCCGTAACAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATTAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATTCCTCTAAACCTGCGCTTTGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGGATCCTCAGAGGCGAGAGGGTCAATTGCTGACTTTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
```

```
Qy 541 CGGCARACGGCTTTCCGCACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGCTGAGAACAGGACGAAAGCTGTTTCATTAACGAGGCAATACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCCGACAATGCTCTAATTAACGGGTGCGGTAAAGCTTCGCAATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGACACAGCAGGAGTCCGGGCGGGGTGGGAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACACAGTGGAAACACCTCTTAAGTTTGATACGCGCAATTTATTTCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACGCGCAGAGTTGCCCGCAGCGCCGACCCCGTTATTGGATTTCGATCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACGCGCCCATCCACCGCAAAACAGCGGTATCGGTGACATTACCACCTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTCCGCGACAGTACTAACTCGCAATCTCGCGCGGACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACGTCTCCGCTGACGCGGATTAACACGCGCGAGTGTGTAACCTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGCTAGCGATACACGACGATGATTCAGTTTCGCTGCTTCGACAGCTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGGCTATRAAACCGCGCTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGCGAAATACGCGAGGCAATGTTGTTGCGCAGGTTTTCACGAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAAGCACGCATACCGCGGTGCGAGTTTGAGATCTCATCACCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

RESULT 5

US-10-601-319-2

; Sequence 2, Application US/10601319

; Publication No. US20040091968A1

GENERAL INFORMATION:

; APPLICANT: Short, Jay M.

; APPLICANT: Kretz, Keith A.

; APPLICANT: Gray, Kevin A.

; APPLICANT: Bazton, Nelson Robert

; APPLICANT: Garrett, James B.

; APPLICANT: O' Donoghue, Eileen

; APPLICANT: Mathur, Eric J.

; TITLE OF INVENTION: RECOMBINANT PHYLASES AND METHODS OF MAKING

; FILE OF INVENTION: AND USING THEM

; FILE REFERENCE: 09010-029011

; CURRENT APPLICATION NUMBER: US/10/601.319

; CURRENT FILING DATE: 2003-06-20

; PRIOR APPLICATION NUMBER: US 09/866,379

; PRIOR FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 09/580,515

; PRIOR FILING DATE: 2000-05-25

```
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-601-319-2
```

Alignment Scores:

Pred. No.:	5,44e-205	Length:	440
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.00%	Indels:	0
DB:	15	Gaps:	0

US-10-034-985-1 (1-1323) x US-10-601-319-2 (1-440)

```
Qy 1 ATGAAGCGATCTTAATCCCATTTTTATCTCTTCTGATCCGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGTTGATTGTCAGTCTGTCATG 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTGTCTCAACCAAGCCACGCACTGATGCAGGATGTCAACCCAGACGCAATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCGGTAAACACTGGGTTGGCTGACACCCGCGNGTGGTGTGAGCTTAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Qy 241 GGCATTTACCAACGCGAGCTCTGTGTAGCCGAGGATGTGCGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGCTCAGCTCCGATTATTGCTGATGTCGACGAGCGGTACCCCTAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCCGCCCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCAGTCCCGATCCGTTATTATTCCTTAAAACTGGCGTTTGGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTGACGCGATCTCTCAGCGGCGAGGAGGTCAATTTGCTGATCTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGCAAAACGGCGTTTCGCAACTGGAACTGGTCTTAATTTTCCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGCTGAGAACAGGACGAAAGCTGTTTCATTAACGAGGCAATACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCCGACAATGCTCTAATTAACGGGTGCGGTAAAGCTTCGCAATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
```

QY 721 GAGATATTTCTCTGCAACAGCACAGGAATCCGCGAGCCGGGTGGGGAAGATCACC 780
 Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACACAGTGGAAACACCTTGCTAAGTTTGATACACGCGCAATTTTATTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 QY 841 CGACGCCAGAGGTGGCCGCGAGCCGCGCCACCCCGCTTATGATTTGATCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTCGCGACACGATACTAATCTGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTCCCGGTACGCGGATTAACACCGCGCAGGTGGTGAACCTGGTTTGAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGCTAAGCGATACACGCGAGTGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGCGTGATAAAACCGCGCTGCTAATAATACCGCGCGCGAGAGGTGAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGCAGGATGGAAGCGGAAATGCGCAGCGCATGTGTTGCTGGCAGGTTTACCCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGATGAAGCACCATACCGCGCTGCGGTGAGTGTGAGATCTCATCACCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 6

US-09-866-379-8
 ; Sequence 8, Application US/09866379
 ; Patent No. US20020136754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KRETZ, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT APPLICATION NUMBER: US/09/866,379
 ; CURRENT FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: US 09/580,515
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 432
 ; TYPE: PRF
 ; ORGANISM: Escherichia coli

US-09-866-379-8

Alignment Scores:
 Pred. No.: 9,266-199 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-866-379-8 (1-432)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTGCTCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCTAGTCTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGCGTGTCTCCACCAAGCCACGCAACTGATGCAGATGTCAACCCAGACGATGCGCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCGCGTAAACACTGGGTTGGCTGCACACCCGCGNGGTGGTGAGCTAATCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGCATTTACCAACGCCAGCGTCTGGTAGCCGCGATTGCTGCGGCGGAAAAAGGCGTCCCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGCTGTGTGAGTCCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAACACAGCGGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTTAATCTCTAAACCTGGCGGTGGTGGCCAACTGGATACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGCTGACTGACCGGATCCCTCAGCAGCGCAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAAACGAGCGATTAACCATCGAACTC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGCTGAGCGCCGACCAATGCTCATTAACCGGTGCGTAAAGCTCGCATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACCAAGCACAGGGAATCCGCGAGCCGGGTGGGGAAGATCACC 780
 Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACACAGTGGAAACACCTTGCTAAGTTTGATACACGCGCAATTTTATTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 QY 841 CGCAGCGCCAGAGGTGCGCGCAGCCCGCCACCCCGTTATTGATTGTTGATCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTACCCACTTCAGTACTG 960


```
QY 1081 CGTCGGCTAAGCGATACACCGCATGCGATTTCAGGTTTCGCTGCTTCCAGACTTACAG 1140
Db |||||||
361 ArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGATAAACCGCGCTCATTAATAACCGCGCGGAGAGGTGAACCTACCC 1200
Db |||||||
381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGGATGTAAGAGCGAAATGCGCAGCGCATGTGTTGCTTGGCAGGTTTACGCAA 1260
Db |||||||
401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGATTG 1296
Db |||||||
421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 8

```
US-10-282-122A-43351
; Sequence 43351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43351
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43351
```

Alignment Scores:

```
Pred. No.: 9,36e-199 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservativeness: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 15 Gaps: 0
```

```
US-10-034-985-1 (1-1323) x US-10-282-122A-43351 (1-432)
QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60
Db |||||||
1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTCTAGTCTGCTATGGT 120
Db |||||||
21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCTGTCTCCAAAGCCAGCGCAACTGATGATCAGAGATGTACCCCGAGCGCATGGCCA 180
Db |||||||
41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCCCGTAAACTGGGTTGGCTGCACACCGCGNGGTGGTGAGCTTAATCGCTATCTC 240
Db |||||||
61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCACGCCAGCGTCTGTGAGCGCGAGTTGTCGCGAAAGAGGCTGCCCG 300
Db |||||||
81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGTGTCAGTCCCGATTATTGCTGATGTCGACGAGGTACCCGTAAACAGCGGAA 360
Db |||||||
101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
Db |||||||
121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATCTCTAAAACTGGCGTTCGCAACTGGATAACGCG 480
Db |||||||
141 SerSerProAspProLeupheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTCACTGACCGCATCTCTCAGCAGGCGAGGGGTCAATTCTGCTGACTTTACCGGCGAT 540
Db |||||||
161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGGCGTTTCGCAACTGGAAACGGGTCTTAATTTCCGCAATCAAACTGTGTC 600
Db |||||||
181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuLys 200
QY 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGCATATTACCATCGGAATC 660
Db |||||||
201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAAGCCGCGACATGCTCTCAATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 720
Db |||||||
221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCCGAGCGCGGGTGGGGAAGGATCACCC 780
Db |||||||
241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTTCACACGAGTGAACACCTTTCCTAAGTTTCATTAACGCGCAATTTTATTGCTACAA 840
Db |||||||
261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
QY 841 CGCAGCCAGAGGTTGCCCGCAGCGCGCGCCACCCGCTTATTGATTTGATCGGACGCG 900
Db |||||||
281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTATCCCACTTCAGTACTG 960
Db |||||||
301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATACATAATCTGGCAAAATCTCGCGCGCGCATGGAGCTCAACTGG 1020
Db |||||||
321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCGCGGTGAGCGGATACACGCGCGAGGTGGTGAACCTGGTGTGAACTGG 1080
Db |||||||
```

Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Qy 1081 CGTCGGCTAAGCATACACCGAGTGTGATTCAGGTTTCGGTCTTCAGACTTACAG 1140
Db 361 ArgGluSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGGTGATAAAGCGCGTGTCTAATAATACGCGCGCGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGAGCGAAATAGCGAGGCGATGTTCGTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 9

US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-8

Alignment Scores:

Pred. No.: 9,26e-199 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 15 Gaps: 0

US-10-034-985-1 (1-1323) x US-10-601-319-8 (1-432)

Qy 1 ATGAACGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCGCATCTGCA 60
Db 1 MetIleAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATTTGTCAGTGTGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40

Qy 121 GTGCGTCTCCAAACCAAGGCCACGCAACTGATCAGGATGTCAACCCAGACGATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCGTAATAACTGGGTTGGCTGCACACCGCGNGGTGGTGAAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Qy 241 GGCATTACCAACGCGCAGCGTCTGGTAGCCGAGGATGCTGGCGGAAAAAGGCGTGGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGCTGGTCCAGTCCGCTATTATTGTGATGTCGACGAGCGTACCCTGTAACAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTGACCGGATCCCTCAGCAGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGCGAGCATTCACATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCCGACATGCTCATTAACCGGTGGGTAGCTCCGATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCACAGGAAATCCGAGCGCGGTGGGAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Qy 781 GATTTCACACGAGTGAACACCTTGTCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACGCGCAGGTTGCGCGCAGCGCGCCGCCCGCTTATTGATTTGATTCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Qy 901 TTCACGCGCCCATCCACGCAAAACAGCGGTATGTTGATACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCGCGCACCATATAATCTGCAAAATCTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACGTTCGCGGTGAGCGGATACACGCGCGCAGGTGGTGAAGTGTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Qy 1081 CGTTCGCTAAGCGATAACACGCGAGTTCAGTTTCGTTGCTTCACAGCTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGCTGATAAAAACCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTCTGTTGGCAGGTTTTTACGCAA 1260


```
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATCAACGACGATACCGGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 10
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003US1
; CURRENT APPLICATION NUMBER: US/10/472,317
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-472-317-41
Alignment Scores:
Pred. No.: 9,26e-199 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 16 Gaps: 0
US-10-034-985-1 (1-1323) x US-10-472-317-41 (1-432)
QY 1 ATCAAGCGATCTATATCCCATTTTATCTCTGATTCGGTTAAACCCCGCAATCGCA 60
Db 1 MetLysAlaIleuIlePheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCACTGCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
QY 121 GTGGTGCTCCAAACGAGCCACGCAACTGATGATGAGGATGTACCCGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCCGTTAAACTGGGTTCGCTGACACCGCGNGGTGTGAGCTTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACGCCAGCGTCTGTAGCCGACGAGATTGTGCGCAAAAAGGGGTGCGCG 300
Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTTCGGGATTTATGCTGATGTGACGAGGTACCCGTTAAACAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGCATCGTTTATTTATCTCTAAAACTCGCGTTTGCCTAACTGATTAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTAGTACGCGATCTTCAGCAGGCGAGGGGTCAATTCGTGACTTTACCGGCAT 540
```

```
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGGCGTTTCGCAACATGGAAACGGGTCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAAACGAGGCATTTACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGTAGGCGCCGACATGCTCTCATTAACGGGTGCGGTAAAGCTCCGATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTCTCAACAAAGACACAGGGAATCCCGAGCCGGGTGGGGAAGATCACCC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGTCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY 841 CGCACCGCAGAGGTTCGCCGACGCGCGCCACCCCGTTATTTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCCCATCCACGCAAAACAGGGGTATGTGTGACATTAATCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGCGACACGATACTAATCTGGCAAAATCTCGGGCGCGACTGAGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGCTTCGGGTGAGCGGATACACGCGCCGAGGTGGTGAACCTGTTGTAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGlyLeuValPheGluArgTrp 360
QY 1081 CGTGGCTAAGCCATAACACGACGATGATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACGCGCTGTCTAATTAATACGCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGGATGTCAAGACGCGAATGCGAGGCGATGTTCGTTGCGCAGGCTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACGACGATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 11
US-10-284-962-3
; Sequence 3, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
```



```
; ORGANISM: Escherichia coli
US-10-284-962-3

Alignment Scores:
Pred. No.:      1.86e-197      Length:      432
Score:          2221.00      Matches:      428
Percent Similarity: 99.07%      Conservative: 0
Best Local Similarity: 99.07%      Mismatches: 4
Query Match:      94.55%      Indels:      0
DB:              15          Gaps:          0

US-10-034-985-1 (1-1323) x US-10-284-962-3 (1-432)

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGATGAGCGGAGCTGAAGCTGAAAGTGTGTGATTCAGTCGTCTGATG 120
Db 21 PheAlaGlnSerGluProGluLeuLeuLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACCAAGGCCAGCAACTGATGCAGGATGTCACCCAGAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Qy 241 GGCATTAACCAACCCAGCGTCTGGTAGCCGACGGATGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGTCAGTCGCGATATTGCTGATCTGCACGAGGTACCGGTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCCCGCGCGCTGGACCTGACTGTGCAATAACCGTACATCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCGATCGTTATTATTAATCTCTCAAACTGGCGTTGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGTGACTACCGGATCTCCTCAGAGCGGAGGCTCAATTGCTGACTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGGTTCGCGAACTCGAACCGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGTGTTCATTAAACGAGGATTAACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCGCAATGCTCTAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGCAACAGCACAGGAAATGCGGAGCGGGTGGGAAGGATCAAC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACCAAGTGAACACCTGCTTAAGTTTGCATAACGCGCAATTTTATTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
Qy 841 CGCACCGCAGAGTTGCCGAGCGCGCACCCCGTTTATGATTGATCATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300

RESULT 12
US-10-284-962-14
; Sequence 14, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Douglas M.
; APPLICANT: Orr, Donald E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-284-962-14

Alignment Scores:
Pred. No.:      1.86e-197      Length:      432
Score:          2221.00      Matches:      428
Percent Similarity: 99.07%      Conservative: 0
Best Local Similarity: 99.07%      Mismatches: 4
Query Match:      94.55%      Indels:      0
DB:              15          Gaps:          0

US-10-034-985-1 (1-1323) x US-10-284-962-14 (1-432)

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGATGAGCGGAGCTGAAGCTGAAAGTGTGTGATTCAGTCGTCTGATG 120
Db 21 PheAlaGlnSerGluProGluLeuLeuLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACCAAGGCCAGCAACTGATGCAGGATGTCACCCAGAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Qy 241 GGCATTAACCAACCCAGCGTCTGGTAGCCGACGGATGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGTCAGTCGCGATATTGCTGATCTGCACGAGGTACCGGTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCCCGCGCGCTGGACCTGACTGTGCAATAACCGTACATCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCGATCGTTATTATTAATCTCTCAAACTGGCGTTGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGTGACTACCGGATCTCCTCAGAGCGGAGGCTCAATTGCTGACTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGGTTCGCGAACTCGAACCGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGTGTTCATTAAACGAGGATTAACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCGCAATGCTCTAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGCAACAGCACAGGAAATGCGGAGCGGGTGGGAAGGATCAAC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACCAAGTGAACACCTGCTTAAGTTTGCATAACGCGCAATTTTATTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
Qy 841 CGCACCGCAGAGTTGCCGAGCGCGCACCCCGTTTATGATTGATCATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
```

```
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTACCAACGCGAGCTCTGGTAGCCAGAGATTGTCGGCAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTCCGGATTATTGCTGATGTCGACGAGGTACCGTAAACACAGCGAA 360
Db 101 GlnProGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATTCCTCTAAACCTGGCGTTGGCCAACTGGATAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTACTCAGCGATCCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAACCGCGGTTCGCGAAGTGAACGGGTGCTTAATTTCCGCAATCAAACTTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
QY 601 CTTAAAGCTGAGAAACAGGACGAAGCTGTTCATTACGCGAGCATTAACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTAGCGCCGACATGCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTGCAACAGCACAGGGAATCCGAGCGCGGTGGGGAAGATCAC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGCTAAGTTTGCAATACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCGCAGAGGTGCGCGCAGCGCCACCGCTTATTGATTTGATTCATCGGACGC 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCCATCCACCGCAAAAACAGCGGTATGTTGATGATTAACCTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGACACGATACTATCTGGCAATCTCGCGCGCGACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGTCAGCCGATAACACCGCGCAGGTGCTGAACCTGGTGTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGTAAACGATACACCGATTCAGTTTCAGTTTCGTTGCTGTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGTATAAACCGCGCTCTCATTAATAACCGCGCGCGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGAGCAAAATGCGCAGGATGTTGTTGCTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAGCACCGATACCGGCGTGAGTTTG 1296
```

```
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 13
US-10-266-041-1
; Sequence 1, Application US/10266041
; Publication No. US2003007284A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-10-266-041-1
Alignment Scores:
Pred. No.: 1,86e-197 Length: 433
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 94.55% Indels: 0
DB: 14 Gaps: 0
US-10-034-985-1 (1-1323) x US-10-266-041-1 (1-433)
QY 1 ATCAAAAGCATCTTAATCCCATTTTATCTCTGATTCCTGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGCAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTGCTGCTCATGGT 120
Db 21 PheAlaGlnSerGlnProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGCTGCTCCACCAACGACGCGCACTGATCAGGATGTCACCCAGACGCGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCGGTAAACACTGGGTTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTACCAACGCGAGCTCTGGTAGCCAGAGATTGTCGGCAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTCCGGATTATTGCTGATGTCGACGAGGTACCGTAAACACAGCGAA 360
Db 101 GlnProGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATTCCTCTAAACCTGGCGTTGGCCAACTGGATAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTACTCAGCGATCCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
```

```
QY 541 CGGCAACGCGGTTTCGCAACGTGAACGCGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
QY 601 CTTAAACGTGAGAACAGGACGAGCTGTTTCATTAACGAGGACGATTAACCATCGGAATC 660
Db 201 LeuAsnArgGluLeuGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGCAACATGCTCATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATGCCGAGCGCGGTGGGAAGATCAACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGACACTTGCTAAGTTTGCATACCGGCNAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGACGCGCAGAGTTGCCGCGCGCCACCGCCGTTATTGGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGATGATPACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCGCGACAGTACTATCTCGCAATCTCGCGCGCGCTGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTCTCCGCTGAGCGGATAACACCGCCGCGTGTGCTGAACCTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGTAAACGCGTGTGATGAGTTCAGTTCGCTGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGGTATAAAACCGCGTGTGATTAATACGCGCGCGCGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTAAGAGCGAAATGCGCAGGCGATGTTGCTTGGCAGGTTTTCAGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 14

```
US-10-284-962-5
; Sequence 5, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 432
; TYPE: PRN
; ORGANISM: Escherichia coli
```

US-10-284-962-5

```
Alignment Scores:
Pred. No.: 3 18e-195 Length: 432
Score: 2197.00 Matches: 424
Percent Similarity: 98.61% Conservativeness: 2
Best Local Similarity: 98.15% Mismatches: 6
Query Match: 93.53% Indels: 0
DB: 15 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-10-284-962-5 (1-432)

```
QY 1 ATGAAGAGCATCTAATCCCATTTTATCTCTTCTGATTCGTTAAACCCCGCATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGAAAGTGTGTGATTGTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCTGCTCCAAACGAGCCAGCAACTGATGCAGGATGTCACCCAGACGCAATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrgly 80
QY 241 GGACATTACCAACGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 101 GlnProGlyGlnValAlaIleIleValAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAACCTGCGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 141 SerSerProAspProLeuPheIleProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGCTGACTCAGCGCATCTCTCAGCAGGCGAGGAGGTCAATTTGCTGCTTATTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACGCGGTTTCGCAACCTGGAACGCGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuAsn 200
QY 601 CTTAAACGTGAGAACAGGACGAGGCTTTCATTAACGAGGACGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAsnGluSerCysAsnLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGCAACATGCTCATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATGCCGAGCGCGGTGGGAAGATCAACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGACACTTGCTAAGTTTGCATACCGGCNAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGACGCGCAGAGTTGCCGCGCGCCACCGCCGTTATTGGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGATGATPACCCACTTCAGTACTG 960
```

```
Db 301 LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACACGATACATACTGGCAAAATCTCGCGGGCCACTGAGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGAGCGGATACACGCCCGCAGGTGGTGAACCTGGTGTGGACCGTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
QY 1081 CGTCCGGTAAGCGATACACCGAGTGGATTTCAGGTTTCGGTGTCTTCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACCGCGCTGTCTCAATTAATACCGCGCCCGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATGTAGAGCGAAATGCCAGGGCATGTGTTTCGTGGCAGGTGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCACGCATACCGCGCTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 15

```
US-10-156-660-2
; Sequence 2, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 2

LENGTH: 436

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: modified phytase enzyme

US-10-156-660-2

Alignment Scores:

Pred. No.:	2,19e-194	Length:	436
Score:	2188.00	Matches:	425
Percent Similarity:	97.70%	Conservative:	0
Best Local Similarity:	97.70%	Mismatches:	10
Query Match:	93.15%	Indels:	0
DB:	14	Gaps:	0

```
US-10-034-985-1 (1-1323) x US-10-156-660-2 (1-436)
QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCGGTTAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGGAGCGCGAGCTGAAGCTGGAAAGTGTGGTGAATTGTCTAGTGGTCTCATG 120
Db 21 PheAlaGlnSerGlnProGlnLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGTCTCAACCAAGCCACGCAATGATGCGAGGATGTCAACCCACGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACTGGCCCGTAAACTGGTGGCTGCACCGCGNGGTGGTGGAGCTAATGCCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACGCGCAGCGTCTGGTAGCCGCGGATTCCTGGCGCAAAAGGGCTGCCCG 300
Db 81 GlyHisTyTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
QY 301 CAGTCTGTGTCAGGTCCGATTATTGCTGATGTCGACGAGCGTACCGCTAAAACAGCGCAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGCTGGCAGCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTCTAAAACTGGCGTTTGGCAACTGGATAACCGG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTCACTGACCGGATCTCTCAGAGGCGAGGGTCAATTGCTCACTTTACCGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTTCGCGAATCTGGAACCGGTCTTAATTTTCGCAATCAAACTGTGTGC 600
Db 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCGATTCACATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAAGCCGCGACATGTTCTCAATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCCGAGCGCGGGTGGGGAAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTTCACCAAGTGAACACCTTGTCTTAAGTTTCGATAACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
QY 841 CGCACGCCAGAGTTGCCCGCAGCGCCACCCCGTTATTGGATTGTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGGCGGTATGGTGTGATGATACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACACGATACATACTGGCAAAATCTCGCGCGGCGACTGGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGAGCGGATACACGCCCGCAGGTGGTGAACCTGGTGTGGACCGTGG 1080
```

```

Db      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Qy      1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
          |||||
Db      361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
          |||||
Qy      1141 CAGATCGGTGATAAAGCGCGCTGTCTAATTAAATACGCCGCCGAGAGGTGAACTGACC 1200
          |||||
Db      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
          |||||
Qy      1201 CTGCGAGGATGTGAAGAGCGAATGCGCAGGCGATGTTCGTTGGCAGGTTTACGCAA 1260
          |||||
Db      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
          |||||
Qy      1261 ATCGTGAATGAAGCACCACATACCGCGCTGCAGTTTGAGATCTCAT 1305
          |||||
Db      421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHis 435
          |||||

```

Search completed: March 12, 2005, 00:05:49
 Job time : 209.5 secs

THIS PAGE BLANK (user #2)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:20:13 ; Search time 45 Seconds
(without alignments)
4389.365 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcacatcaccatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10034985/runat 10032005_164241_11973/app_query.fasta_1.1479
-DB=issued Patents AA -OFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985 @CGN 1.1 46 @runat 10032005_164241_11973 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	98.0	440	3	US-09-259-214-2
2	2302	98.0	440	3	US-09-318-528-2
3	2302	98.0	440	3	US-09-231-931-2
4	2302	98.0	440	4	US-09-580-515-2
5	2221	94.6	433	4	US-09-540-149A-1
6	2194.5	93.4	423	2	US-08-910-798-2
7	528	22.5	421	4	US-09-489-039A-7512
8	376	16.0	522	4	US-09-489-039A-13501
9	119	5.1	439	3	US-09-044-718-3
10	119	5.1	439	4	US-10-062-848-3
11	119	5.1	440	4	US-09-684-855-106
12	119	5.1	440	4	US-09-684-855-128

13	119	5.1	440	4	US-09-684-855-151	Sequence 151, Appl	
14	119	5.1	440	4	US-09-488-265B-6	Sequence 6, Appli	
15	119	5.1	449	3	US-09-044-718-12	Sequence 12, Appl	
16	119	5.1	449	3	US-10-062-848-12	Sequence 12, Appl	
17	119	5.1	465	3	US-08-868-435-33	Sequence 33, Appl	
18	119	5.1	465	3	US-08-744-231-33	Sequence 33, Appl	
19	119	5.1	465	3	US-09-044-718-78	Sequence 78, Appl	
20	119	5.1	465	4	US-09-636-499-6	Sequence 6, Appli	
21	119	5.1	465	4	US-09-273-871A-8	Sequence 8, Appli	
22	119	5.1	465	4	US-10-083-452-8	Sequence 8, Appli	
23	119	5.1	465	4	US-09-635-504-33	Sequence 33, Appl	
24	119	5.1	465	4	US-10-062-848-78	Sequence 78, Appl	
C	25	118.5	5.1	575	4	US-09-252-991A-20649	Sequence 20649, App
26	118	5.0	440	4	US-09-684-855-107	Sequence 107, App	
27	118	5.0	440	4	US-09-684-855-129	Sequence 129, App	
28	118	5.0	440	4	US-09-684-855-152	Sequence 152, App	
29	118	5.0	440	4	US-09-488-265B-7	Sequence 7, Appli	
30	117	5.0	386	1	US-08-758-213-1	Sequence 1, Appli	
31	117	5.0	386	2	US-08-692-787-48	Sequence 48, Appli	
32	117	5.0	386	3	US-09-097-199-48	Sequence 48, Appl	
33	117	5.0	386	4	US-09-949-016-6022	Sequence 6022, Ap	
34	117	5.0	515	2	US-09-146-283-2	Sequence 2, Appli	
35	117	5.0	515	3	US-08-579-823A-2	Sequence 2, Appli	
36	117	5.0	515	3	US-09-344-195-2	Sequence 2, Appli	
C	37	116.5	5.0	330	4	US-09-252-991A-32186	Sequence 32186, A
38	116.5	5.0	442	4	US-09-949-016-9351	Sequence 9351, Ap	
39	116.5	5.0	442	4	US-09-949-016-9352	Sequence 9352, Ap	
40	116.5	5.0	442	4	US-09-949-016-9353	Sequence 9353, Ap	
41	116.5	5.0	442	4	US-09-949-016-9354	Sequence 9354, Ap	
42	116.5	5.0	442	4	US-09-949-016-9355	Sequence 9355, Ap	
43	116.5	5.0	450	4	US-09-765-111A-14	Sequence 14, Appl	
C	44	116.5	5.0	1093	4	US-09-252-991A-21827	Sequence 21827, A
C	45	115.5	5.0	539	4	US-09-252-991A-33061	Sequence 33061, A

ALIGNMENTS

RESULT 1
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259, 214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910, 798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Alignment Scores:
Pred. No.: 9.08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-259-214-2 (1-440)

Qy	1	ATGAAGCGCATCTTAATCCCATTTTATCTCTCTCGTTAACC	CGCAATCGCA 60
Db	1	MeLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla	20
Qy	61	TTCCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGATTGTCAGTCGTGCT	120

```
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGGTCTCAACCAAGGCCAGCACTGATGAGGATGTCACCCAGCAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTGGCTGACACCGCGNGGGTGGTGAAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACGCGGTCTGGTAGCCAGCGATTGCTGCGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGCTGGTCAAGTCGGATTATTTGCTGATGTCGACAGCGTACCCGTAACCAAGCGCAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGGCTGGCCACTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTTAATCTCTAAACTGGCGTTTGGCAACTGGATAACGG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTCACGCGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAAGCTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGCATATACCATCGAACTC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCCCGCACATGCTCATTAACCGGTGGGTAGCTCGCATCAATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGAATCCGAGCGCGGGTGGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGACACTTGTCAAGTTTGCATACCGCAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCGAGGTTGCGCGCAGCGCGCCACCCCGTTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCGCCATCCACGCGAAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGCGACAGTACTAATCTCGCAATCTCGCGCGCGCAGTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACCTTCCCGGTGAGCGGATAACACGCGCGAGGTGGTGAACCTGGTGTGAAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTGGCTAAGCGATAACAGCCAGTGGATTCAAGTTTCGTGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGCTATAAAACCGCGTGTCAATTAATACGCGCGCGCGAGGAGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
```

```
QY 1201 CTGGCAGGATGTGACAGCGAAATGCGAGGCGATGTGTTGTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCAGCAGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCATCAC 1320
Db 421 IleValAsnGluAlaAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440

RESULT 2
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

Alignment Scores:
Pred. No.: 9,08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-318-528-2 (1-440)
```

```
QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGGTGTGATGTCAGTCGTCATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGCGTCTCCAAACCAAGGCCAGCAACTGATCGAGATGTCAACCCAGCAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTGGCTGACACCGCGNGGGTGGTGAAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACGCGGTCTGGTAGCCAGCGATTGCTGCGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGCTGGTCAAGTCGGATTATTTGCTGATGTCGACAGCGTACCCGTAACCAAGCGCAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGGCTGGCCACTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTTTGGCAACTGGATAACGG 480
```



```
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTACCGATCTCTCAGCAGGCGAGGGTCAATGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTTCGGCACTCGAACGGTGTCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGCAGCAAGAGTGTCTTAAACGAGGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCCGCGCAATGCTCTAATTAACCGGTGCGGTAAAGCTTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTTCACACAGTGGAAACCTTGCTAAGTTTGCAATACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCAGCCAGAGGTGCCCCAGCGCCGCCACCCGTTATTGGATTGATCATGCGACG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCGGACACGATCTAATCTGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACCTTTCCCGGTGAGCGGATAACACCGCCGAGGTGGTGAACCTGGTGTGTTGAACGTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGCTAAGCATATACACCGCATGGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGCGTATAAACCGCGCTGTCTAATTAATACCGCCCGCGAGAGGTGAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGGCAGGATGTGAAGACGAAATGCGCAGGCAATGTGTGTTGTCGAGGTTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAAGACGCATACCGCGGTGAGTTTGAGATCTCATACCATCATCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

RESULT 3

```
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0.
```

```
; SEQ ID NO 2
```

```
; LENGTH: 440
```

```
; TYPE: PRT
```

```
; ORGANISM: Escherichia coli
```

```
US-09-291-931-2
```

Alignment Scores:

```
Pred. No.: 9,08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 3 Gaps: 0
```

```
US-10-034-985-1 (1-1323) x US-09-291-931-2 (1-440)
```

```
Qy 1 ATCAAAAGCATCTTAATCCCATTTTATCTCTCTGATTCGTTAACCCCGCAATCTGCA 60
```

```
Db 1 MeClySAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 20
```

```
Qy 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGCTGATTTGTCACTCGTCATGGT 120
```

```
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
```

```
Qy 121 GTGCTGTCTCCAAACCAAGGCCACGCAACTGATGTCAGGATGTCACCCAGACGATGGCCA 180
```

```
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
```

```
Qy 181 ACCTGCGCGGTAAACCTGGTGGCTGACACCGCGNGGTGGTGGCTAATCGCTATCTC 240
```

```
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
```

```
Qy 241 GGACATTACCAACCGCAGCGGTCTGTGTAGCCGACGAGTTGTGCGGAAAAAGGGCTGCCG 300
```

```
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
```

```
Qy 301 CAGTCTGGTCAGTCCGATATTGCTGATGTCGACGAGCGTACCCGTAAAAACAGGCGAA 360
```

```
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
```

```
Qy 361 GCCTTCGCGCGCGGTGGCCACTGCTGCAATACCGTACATACCCAGGAGGATACG 420
```

```
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
```

```
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAATCTGGCGTTTCCCACTGGATAACGCG 480
```

```
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
```

```
Qy 481 AACGTGACTGACCGCATCTCTACGAGGCGAGGAGGTCAATGCTGACTTTTACCGGCAT 540
```

```
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
```

```
Qy 541 CGGCAACCGCGTTTCGGCACTCGAACCGGTGCTTAATTTCCGCAATCAAACTGTGC 600
```

```
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
```

```
Qy 601 CTTAAACGTGAGAAACAGCAGCAAGCTGTTCATTAAACGAGGATTAACCATCGGAATC 660
```

```
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```

```
Qy 661 AAGGTGAGCGCGCAATGCTCTAATTAACCGGTGCGGTAAAGCTTCGCATCAATGCTGACG 720
```

```
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
```

```
Qy 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGGATCAC 780
```

```
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
```

```
Qy 781 GATTTCACACAGTGGAAACCTTGCTAAGTTTGCAATACCGCAATTTTATTTGCTACAA 840
```

```
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
```

```
QY 841 CGCAGCCGAGGTGTCGCCGAGCGGCCACCCCGTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGTGATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACACGATACTTAATCTGGCAAAATCTCGCGCGGCACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGACCGGATAACACCGCCCGAGTGTGAACCTGTTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGGTAAAGCAGTAAACCCAGTGTGATTCAGGTTTCGGTGTTCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGATAAAACCGCGCTGTCTTAATAATACGCCCGCGAGAGGTGAAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGCAGCGCATGTGTTTCGTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATGCTGAATAGCAGCAGCATACCGGCTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

RESULT 4

```
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-580-515-2
```

```
Alignment Scores:
Pred. No.: 9,08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 4 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-09-580-515-2 (1-440)

```
QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTCTGATTCCTGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTAGCCGGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTCGTCAATGGT 120
```

```
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCCGTGCTCCACACCGCCAGCCCAACTGATCAGGATGTACCCAGCAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCCCGTAAACCTGGGTTGGCTGACACCCCGNGGTGCTGAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTATCAACCCAGCGGTCTGGTAGCCGCGATTCTCTGCGGAAAAAGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAAGTCCGGATTATTGCTGATGTCGACGAGCGTACCCGTAACAGAGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGCGACCTGACTGTGCAATAACCGTACATACCCAGCAGCATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATCTCTAAAACTGGCGTTTGGCCAACTGGATACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGCTCAGTCAGCGATCTCTCAGCAGGCGAGGAGTCAATTGCTGACTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAAAACGGCGTTTCGCAACTGGAAACGGGTCTTAAATTTCCGCAATCAAACTTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAAACAGACGAAAGCTGTTTCAATTAAACGAGCATTAACATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTTCAGCGCCGACAAATGCTCATTAACCGGTGCGTAAGCCCTGCATCAATCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTCGCAACAAAGCACAGGGAATCCCGAGCGCGGTGGGGAAGATCAC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTCAACAGTGGAAACACTTGTAGTTTCATTAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCCGAGAGGTTCGCCGAGCGCGCCACCCCGTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGTGATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACACGATACTTAATCTGGCAAAATCTCGCGCGGCACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGACCGGATAACACCGCCCGAGTGTGAACCTGTTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGGTAAAGCAGTAAACCCAGTGTGATTCAGGTTTCGGTGTTCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGATAAAACCGCGCTGTCTTAATAATACGCCCGCGAGAGGTGAAACTGACC 1200
```

Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGAGGATGTGAAGAGCGAAATGCCAGGCGATGTTCTGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATCAGACGACGATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 5
US-09-540-149A-1
; Sequence 1, Application US/09540149A.
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Kingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; PRIORITY FILING DATE: 2000-03-31
; PRIOR FILING DATE: 60/127,032
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Alignment Scores:

Prod. No.: 2,26e-221 Length: 433
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 94.55% Indels: 0
DB: 4 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-540-149A-1 (1-433)

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAACCCTGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGTGATGTCAGTCGTATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACACGAGCGCAGCACTGATCAGGATGTCACCCAGACGATGCGCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCGGTAAACCTGGGTGTGCTGACACCGCGGTGTGAGCTAACTCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Qy 241 GGACATTAACACCGCAGCGTCTGTTAGCCGACGATGTCGGCGAAAGGCGTCCCG 300
Db 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGTCAGTCCGATTTATGCTGATGTCGACGAGCGTACCGTAAACAGCGGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAAATCGCGGTGTCCTGCAACGCG 480

Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGCTGACTCAGCCGATCTCTCAGCAGGCGAGGAGGTCAATTTGCTGACTTTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGCCTAAACGGCGTTTCGCAACTGGAAACGGGTCTTAATTTTCCGCNATCAAACTGTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTAAACGTCAGAAACAGGACGAAAGCTGTTCAATTAACGCGAGCATTAACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCCGACCAATGTCTCATTAACCGGTGCGGTAAAGCTTCGTCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCAGGGAATCCGAGCGCGGGTGGGGAAGATCAAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACGAGTGAACACCTTGTAACTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Qy 841 CGCACCGCAGAGGTTCGCGCAGCGCCGACCCGTTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACCCCGCATCCACCGCAAAACAGGCGTATGTTGTGATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCGGACGACGATCTATCTGCAAAATCTCGCGCGCAGCTGAGGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACGTTTCCCGGTGAGCGGATTAACCGCGCAGGCGTGTGTAACCTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTGGCTAAGCAATTAACGCGAGTGGATTCAGTTTCGCTGCTTCTTCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATCGGTGATAAAGCGCGCTGCTCATTAATAACGCCCGCGAGAGGTGAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGAGGATGTGAAGAGCGAAATGCGAGGCGATGTTGTTGTCGAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAACGACGATACCGCGGTGCGAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 6

US-08-910-798-2
; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US

```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,798
; FILING DATE: August 13,1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HALL, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-910-798-2

```

```

Alignment Scores:
Pred. No.: 1,25e-218 Length: 423
Score: 2194.50 Matches: 423
Percent Similarity: 96.14% Conservative: 0
Best Local Similarity: 96.14% Mismatches: 0
Query Match: 93.42% Indels: 17
DB: 2 Gaps: 1

```

US-10-034-985-1 (1-1323) x US-08-910-798-2 (1-423)

```

QY 1 ATGAACGGATCTTAATCCCATTTTATCTCTGATTCGGTTAACCCTCGCATCTGCA 60
DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAAGCTGGAAGTGTGGTGTGATTCAGTTCAGTTCAGTTCAGT 120
DB 21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGGTCTCAACCAAGGCCAGCAACTGATGAGGATCTACCCAGCAGCATGCCCA 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCCGTAAACTGGTGTGGCTGACACCGCGGCGGTGGTGGATTAATCGCTATCTC 240
DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTACCAACGCCAGCTGCTGGTAGCCGACGGATTGCTGGCAAAAAGGGCTGCCCG 300
DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTTCGGATTATTGCTGATGTGACAGAGCGTACCCGTAACACAGCGAA 360
DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGly--- 119
QY 361 GCCTTCGCCCGCGGTGGACCTGACTGTGCAATAACCGTACATACCCAGCAGATACG 420
DB 120 -----GlnAlaAspThr 123
QY 421 TCAGTCCCGATCCGTTATTATTAATCTCTTAACCTGGCGTTTGCCCACTGGATAACGCG 480
DB 124 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 143
QY 481 AAGTGTGACTGACGGATCTCAGCAGGCGAGGAGGTCATTCGTGACTTTACCGGCAT 540
DB 144 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 163
QY 541 CGCAACAGCGGTTCGCAACTGGACGGGTGCTTAATTTCCGCAATCAAACTGTGC 600

```

```

DB 164 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 183
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGCAGCATTAACCATCGAACTC 660
DB 184 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 203
QY 661 AAGTGTGAGCGCCGACAAATGTTCTCAATTAACCGGTGCGGTAAAGCTCGCATCAATTCCTGACG 720
DB 204 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 223
QY 721 GAGATATTCTCTCTGCAACAGCAGCAGGAGGATCCGAGCGCGGGTGGGGAAGATCACC 780
DB 224 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 243
QY 781 GATTTCACACGAGTGAACACCTTTGCTAAAGTTTCATTAACCGGTGCGGTAAAGCTCGCATCAATTCCTGACAA 840
DB 244 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 263
QY 841 CGCACGCCAGAGTTGCCCGCAGCCGCGCACCCCGTTATTGATTTGATCATGGCAGCG 900
DB 264 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 283
QY 901 TTGACGCCCATCCACGCGCAAAACAGGCGGTATGTTGTGACATTACCCACTTCAGTACTG 960
DB 284 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 303
QY 961 TTTATTGCGCGACACGATATACTATCTGGCAAAATCTCGCGCGGCACCTGGAGCTCAACTGG 1020
DB 304 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 323
QY 1021 ACGTTTCCCGTCCAGCCGAGTAAACAGCCGCGCAGGTGGTGAACCTGTTTGAACGCTGG 1080
DB 324 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 343
QY 1081 CGTCCGCTAAGCATTAACAGCCAGTTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
DB 344 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 363
QY 1141 CAGATCGGTGATAAAACGCGCTGTCTATTAAATACGCCGCCGAGAGGTGAAACTGACC 1200
DB 364 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 383
QY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGCAGGGGATGTGTTGTTGGCAGGTTTACGCAA 1260
DB 384 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 403
QY 1261 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTCAGATCTCATCACCATCACCATCAC 1320
DB 404 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 423

```

RESULT 7

```

US-09-489-039A-7512
; Sequence 7512, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Braton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7512
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7512

```

```

Alignment Scores:
Pred. No.: 7,89e-46 Length: 421
Score: 528.00 Matches: 134

```

Db 324 AlaserLeuLeuThralaLeuAaspPhelysProTyrGlnLeuProGlyGlnTyrGluArg 343

Qy 1045 ACQCCGCCAGGTGGTGAACCTGGTGTGGTAAACGCTGGCTGGCTAAACGATACAGCCAG 1104

Db 344 ThrProileidGlyLysLeuLeuPheGlnArgTrpHisAaspSerAlaGlyAenArgAasp 363

Qy 1105 TGGATTACAGTTTCGCTGGTCTCCAGACTTTACAGCAGATGGCTGTATATAAACCCCGCTG 1164

Db 364 LeuMetLysIleGluTyrValTyrGlnSerThrGlnLeuArgAenAlaAaspAlaLeu 383

Qy 1165 TCATTAAATACGCCGCCGAGAGAGTCAAACTGACCTGGCAGCATGTGAAGAGCGAAAT 1224

Db 384 ThrLeuGlnAlaProProGlnArgValThrLeuAlaLeuAenGlyCysProVal--Aasp 402

Qy 1225 GCCAGAGCCATGTTCGTTGGCAGGTTTACGCAAAATCGTGAATGAAGCA 1275

Db 403 AspGlnGlyPheCysProLeuGluThrPheLysMetValIleAenGluAla 419

RESULT 8

US-09-489-039A-13501

; Sequence 13501, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 13501

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13501

Alignment Scores:

Pred. No.: 5,08e-30 Length: 522

Score: 376.00 Matches: 115

Percent Similarity: 42.70% Conservative: 75

Best Local Similarity: 25.84% Mismatches: 209

Query Match: 16.01% Indels: 46

DB: 4 Gaps: 15

US-10-034-985-1 (1-1323) x US-09-489-039A-13501 (1-522)

Qy 4 AAGCGGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCCG---CAATCTGCA 60

Db 107 GlnGlyLeuLeuArgLeuPheIleAlaCysAlaLeuProLeuAlaLeuGlnSerAla 126

Qy 61 TTGCTCAGAGTGAGCCGGAGCTGAACCTGAAAGTGTGGTGAATTGTCAGTCGTCATCGT 120

Db 127 AlaAla-----AlaAaspTrpGlnLeuGluLysValValGluLeuSerArgHisGly 143

Qy 121 GTGCGTGCTCCAAACGAGCCACG---CAACTGATGCAGGATGTCACCCACAGCGCATGG 177

Db 144 IleArgProThrAlaGlyAenArgGluAlaIleGluAlaAlaThrGlyArgProTrp 163

Qy 178 CCAACCTGGCGGTAAACCTGGTGTGCTGCTGACACCGCGNGGTGTGTAGCTAATCGCTAT 237

Db 164 ThrGluTrpThrThrHisAaspGlyGluLeuThrGlyHisGlyTyrAlaAlaValAen 183

Qy 238 CTCGGACATTACCAACGCCAGCGCTCGTAGCCGACGATTGCTGGGGAAGAGGCTGC 297

Db 184 LysGlyArgAlaGluGlyGlnHisTyrArgGlnLeuGlyLeuLeu---GlnAlaGlyCys 202

Qy 298 CCGCAGTCTGTGTCAGGTCGGGATTATTGCTGATGTCGACGAGCGTACCCGTTAAACAGGC 357

Db 203 ProThrAlaGluSerIleTyrValArgAlaSerProLeuGlnArgThrArgAlaThrAla 222

Qy 358 GAAGCCTTCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCAGGAGAT 417

```
Db 223 GlnAlaLeuValAspGlyAlaPheProGlyCysGlyValAlaIleHisTyrVal----- 240
Qy 418 ACCTCCAGTCCGATCCGTTATTATTCCTCTAAACTCGGTTCCCAACTGGATAAC 477
Db 241 SerGlyAspAlaAspProLeuPheGlnThrAspLysPheAlaAlaThrGlnThrAspPro 260
Qy 478 GCGAAGCTGACTGACGGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGG 537
Db 261 AlaArgGlnLeuAlaAlaValLysGluLysAlaGly-----AspLeuAlaGln 276
Qy 538 CATCGCAAAACGGCGTTTCGCGAATCGAAACGGGTGCTTAATTTTCGCAATCAAACTTG 597
Db 277 ArgArgGlnAla-----LeuAlaProThrIleGlnLeuLeuLysGlnAlaVal 292
Qy 598 TGCCTTAACGTGAGAAACAGGAGCAAGCTGTTCAATTAACGAGGATTAACATCGGAA 657
Db 293 Cys-----GlnAlaAspLysProCysProIlePheAspThr---ProTrpGln 307
Qy 658 CTCAGGTGAGCGCGACATGCTCATTAACCGGTGCGGTAGCCTCGCATCAATGCTG 717
Db 308 ValGluGlnSerLysSerGlyLysThrThrIleSerGlyLeuSerValMetAlaAsnMet 327
Qy 718 ACGAGATATTCTCTCGCAACAGCACAGGGAATGCCG-----GAGCGGGGTGGGA 771
Db 328 ValGluThrLeuArgLeuGlyTrpSerGluAsnLeuProLeuSerGlnLeuAlaTrpGly 347
Qy 772 AGGATACCGATTCACACCAAGTGAACACCTTGCTGAAGTTTGATTAACGCGCAATTTAT 831
Db 348 LysIleThrGlnAlaArgGlnIleThrAlaLeuLeuProLeuLeuThrGluAsnTyrAsp 367
Qy 832 TTGCTACAACGCGCCAGAGGTGCGCGAGCGCGCCACCCGCTATTGGATTTGATC 891
Db 368 LeuThrAsnAspValLeuTyrThrAlaGlnLysArgGlySerValLeuLeuAsnAlaMet 387
Qy 892 ATGCGACGTGTAGCGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTACCCACT 951
Db 388 LeuAspGlyValLysProGluAlaAsnProAsnValArgTrp----- 401
Qy 952 TCAGTACTGTTATTGCGGACAGCATACTAATCTGGCAATCTCGCGCGGCACTGGAG 1011
Db 402 -----LeuLeuLeuValAlaHisAspThrAsnIleAlaMetValArgThrLeuMetAsn 419
Qy 1012 CTCAACTGGACGCTTCCCGT---CAGCGGATAACACGCGCGCAGTGTGTAAGTGTG 1068
Db 420 PheSerTrpGlnLeuProGlyTyrSerArgGlyAsnIleProProGlySerSerLeuVal 439
Qy 1069 TTTGAACGCTGGCTCGGCTAAGCGATPAACAGCGCATGAGTTCAGGTTTCGCTGCTTC 1128
Db 440 LeuGluArgTrpArgAsnAlaLysSerGlyGluArgTyrLeuArgValTyrPheGlnAla 459
Qy 1129 CAGACTTTACGAGATGCGTGATAAAACCGCGCTGTCTAATTAATAGCGCGCGGAGAG 1188
Db 460 GlnGlyLeuAspAspLeuArg-----ArgLeuGlnThrProAspAlaGln 474
Qy 1189 GTGAAACTGACC-----CTGCGAGATGTGAAGCGAATGCGCAGGCG 1233
Db 475 HisProMetLeuArgGlnGluTrpHisGlnProGlyCysArgGlnThrAspValGlyThr 494
Qy 1234 ATGTGT-----TCGTTGGCAGGTTTACGCAATCTGTAATGAAGCAGCGATA 1281
Db 495 LeuCysProPheGlnAlaAlaIleThrAlaLeuGlyGlnArgIleAspArgSerAla 514
Qy 1282 CCGCGTGCAGTTG 1296
Db 515 ProAlaValAlaMet 519
```

RESULT 9

US-09-044-718-3

; Sequence 3, Application US/09044718

; Patent No. 6391605

; GENERAL INFORMATION:

; APPLICANT: KOSTREWA, Dirk

```
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
; US-09-044-718-3

Alignment Scores:
Pred. No.: 0.00207 Length: 439
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 3 Gaps: 24
```

US-10-034-985-1 (1-1323) x US-09-044-718-3 (1-439)

```
Qy 19 CCATTTTATCTCTCTCTGATTCGTTAAACCCCAATCTGCATTCGTCAGAGTGAGCCG 78
Db 26 ProPheSerLeuGluAspGluLeuSerValSerSerLysLeuProLys----- 42
Qy 79 GAGCTGAAGCTGGAAGTAGTGTGTGATTGTCAAGTCGTATGCTGTGCTGTGCTCAACC--- 135
Db 43 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer 62
Qy 136 -----AAGCCACGCAACTGATGCAGGATGTACCCCGACAGCA----- 174
Db 63 SerLysSerLysLysTyrLysLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 82
Qy 175 -----TGCCCAACTGCGCGGTAAACTGAGTGGTGG-----CTGACA 210
Db 83 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 102
Qy 211 CCCTGGTGTGCTGAGTAAATCGCTATCTCGGA-----CATTACCAACCCAGCGTCTG 264
Db 103 ProPheGlyGluGlnGlnLeuValAsnSerGlyLysPheTyrGlnArgTyrLysAla 122
Qy 265 GTAGCCGACGATGTCTGGCGAAAGGCGTCCCGCAGTCTGTGTCAGGTCGCGATTATT 324
Db 123 LeuAlaArgSerValPro-----PheIleArg 132
Qy 325 GCTGATGTGACGAGCGTATCCCGTAAACAGCGGAGACCTTCGCGCGCGGCTG----- 378
Db 133 AlaSerGlySerAspArgValIleAlaSerGlyGluLysPheIleGluGlyPheGlnGln 152
Qy 379 -----GCACCTGACTGTGCAATAACCGTA 402
Db 153 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 172
Qy 403 CATACCCAGCAGATACGTCACGTCGATCCGCTTATTTAATCCTCTTAAACCTGGCGTT 462
Db 173 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 186
Qy 463 TGCCAACTGGATAACGCGAACGTCGACTGACGCGATCCTCAGCAGGCGAGGCGTCAATT 522
Db 187 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 199
Qy 523 ---GCTGACTTTACGGGCGATCGGCAACGGCGTTTCGGGAACCTGGACGGGTGCTTAAT 579
Db 200 AlaAlaAsnPheThr-----AlaLeuPhe 207
```

```
QY 580 TTTCCGCAATCAACTTGTGCTTAAACGTCGAGAAACAGGAGCAAGCTGCTTCAATACG 639
Db 208 AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 225
QY 640 CAGGCATTACCATCGGAACCTC-----AAGGTGAGCGCCGACCAATGCTCATTAACCGGT 693
Db 226 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaArgThrSer 245
QY 694 GCGGTAAAGCTCGCATCAATGTCAGCGAGATATTTCTCTGCAACAGACAGGGAATG 753
Db 246 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 261
QY 754 CCGGAGCCGGGTGGGAAGGATCACCAGTTCACACAGTGGACACCTTGCTTAAGTTTG 813
Db 262 -----Trp-----LysLysTyraSerLeuGlnSerLeu 271
QY 814 CATACCGCGCAATTTTATTTGTTACACGACGCGAGGTGCGCGACGCGCGCAC 873
Db 272 -----GlyLysTyraSerLeuGlnSerLeuGlnSerLeuGlnSerLeuGlnSerLeu 289
QY 874 CCGTTATTTGATTTGATCGGAGCGCTTACGCGCCCATCCAGCGCAAAACAG----- 927
Db 290 GlyPheThrAsnGluLeuAlaArgLeuThrArgSerProValGlnAspHisThrSer 309
QY 928 -----GCGTATGTTGTCACATTACCCACTTCAGTACTGTTTATTGCC 969
Db 310 ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyraVal 329
QY 970 -----GGACAGCATATACTCGCAATCTCGCGCGGCGACTGGAGCTCAACTGGACG 1023
Db 330 AspPheSerHisAspAsnSerMetValSerIlePheAlaLeu----- 344
QY 1024 CTTCCCGGTGAGCGGATAACACGCGCCGAGGTGGTGAACCTGGTGTGAACTGGCGGT 1083
Db 345 -----GlyLeuTyraSerLeuGlnSerLeuGlnSerLeuGlnSerLeuGlnSerLeu 1131
QY 1084 CCGCTAAGCGAT---AACAGCAGTGGAT---CAGGTTTCGCTGCTTCCAG 1131
Db 363 GluLeuAspGlyTyraSerAlaSerTrpValValProPheGlyAlaArgAlaTyraPheGlu 382
QY 1132 ACTTTACAGCAGATGCTGATAAACCGCGCTGCTATTAATACGCGCGCGGAGAGGTG 1191
Db 383 ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 401
QY 1192 AAACCTGACCTGGCAGGATGTGAAGCGAAATCGCGAGGCGATGTTGCTGGCAGGT 1251
Db 402 ValValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 420
QY 1252 TTACGCAATCGTGAATGAAGCAGC 1278
Db 421 PheValLysGlyLeuSerTrpAlaArg 429
```

RESULT 10

```
US-10-062-848-3
; Sequence 3, Application US/10062848
; Patent No. 6734004
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/10/062,848
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-062-848-3

Alignment Scores:
Pred. No.: 0.00207 Length: 439
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 4 Gaps: 24
```

US-10-034-985-1 (1-1323) x US-10-062-848-3 (1-439)

```
QY 19 CCATTTTATCTCTCTCGATTTCGTTAAACCCCGCAATCTGCAATTCGCTCAGATGAGCCG 78
Db 26 ProPhePheSerLeuGluAspGluLeuSerValSerLysLeuProLys----- 42
QY 79 GAGCTGAAGCTGGAAGTGTGTGATTGTGTCAGTCGTGTCGTGTCGTGTCCTCAAC-- 135
Db 43 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyraProThrSer 62
QY 136 -----AAGCCACGCAACTGATGTCAGGATGTCACCCGACGCA----- 174
Db 63 SerLysSerLysLysTyraSerLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 82
QY 175 -----TGCCCAACTGCGCGTAAACCTGGTGG-----CTGACA 210
Db 83 LysGlyLysPheAlaPheLeuLysThrTyraSerLysLeuGlyAlaAspAspLeuThr 102
QY 211 CCCGNGGTGTCGATGATGTCGCTATCTCGA-----CATTACCAACGCCAGCGCTG 264
Db 103 ProPheGlyGlnGlnLeuValAsnSerGlyLysPheTyraGlnArgTyraLysAla 122
QY 265 GTAGCCGACGATGTCGTGGGAAAAAGGCGTCCCGCAGTCTGTCAGTCGCGGATATT 324
Db 123 LeuAlaArgSerValValPro-----PheIleArg 132
QY 325 GCTGATGTCGACGAGGTACCCGTAACAGCGGAAGCTTCCGCGCGGCGTG----- 378
Db 133 AlaSerLysArgAspArgValIleAlaSerGlyGluLysPheIleGluGlyPheGln 152
QY 379 -----GCACCTGACTGTGCAATAACCGTA 402
Db 153 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 172
QY 403 CATACCCAGCAGATAGTCCTCCGATCCCGATCCCTTATTTATCTCTTAAACATGGCGTT 462
Db 173 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 186
QY 463 TGCCAACTGGATAACGGAACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCAATT 522
Db 187 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 199
QY 523 ---GCTGACTTTACCGGCGATCGGCAACCGCGTTCGCGCAACTGGAACGGGTCTTAAT 579
Db 200 AlaAlaAsnPheThr-----AlaLeuPhe 207
QY 580 TTTCCGCAATCAAACTTTGTCCTTAAACGTGAGAAACAGACAGCAAGCTGCTTCAATACG 639
Db 208 AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 225
QY 640 CAGGCATTACCATCGGAACCTC-----AAGGTGAGCGCGCGACCAATGCTCTATTAACCGGT 693
Db 226 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaArgThrSer 245
QY 694 GCGGTAAAGCTCGCATCAATGTCAGCGAGATATTTCTCTGCAACAGACAGGGAATG 753
Db 246 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 261
QY 754 CCGGAGCCGGGTGGGAAGGATCACCAGTTCACACAGTGGACACCTTGCTTAAGTTTG 813
```

```
Db      262  -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 271
QY      814  CATAACCGCGCAATTTATTGTTGTTACAAAGCAGCGAGAGTTGCCCGCAGCGCGGCCACC 873
Db      272  -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 289
QY      874  CCGTTATTGATTGATCATGGCAGCGTTGACGCCCGCCCATCCACCGCAAAACAG----- 927
Db      290  GlyPheThrAsnGlnLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 309
QY      928  -----GGGTATGTTGTTGACATTAACCCACTTCCAGTACTGTTTATTGGC 969
Db      310  ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 329
QY      970  -----GGACACGATATACTCGGCAATCTCGCGCGCGCATCGAGCTCAACTGGACG 1023
Db      330  AspPheSerHisAspAsnSerMetValSerIlePheAlaLeu----- 344
QY      1024  CTTCCCGGTGAGCGGATAACAGCGCGCCAGGTGGTGAACCTGGTGTGAACTGGCGT 1083
Db      345  -----GlyLeuTyrAsnGlyThrGluProLeuSerArgThrSerValGluSerAlaLys 362
QY      1084  CGCTTAAGCGAT---AACACCGAGTGGATT-----CAGGTTTCGCTGCTTCCAG 1131
Db      363  GluLeuAspGlyTyrSerAlaSerTrpValValProPheGlyAlaArgAlaTyrPheGlu 382
QY      1132  ACTTACAGCAGATGCGTGATAAAACGCCGCTGTCTATTAATACGCGCGCGGAGAGGTG 1191
Db      383  ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 401
QY      1192  AAAGTACCTGCGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGCTGTTGTTGGCAGGT 1251
Db      402  ValValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 420
QY      1252  TTTACGCAATCGTGAATGAAGCACGC 1278
Db      421  PheValLysGlyLeuSerTrpAlaArg 429

RESULT 11
US-09-684-855-106
; Sequence 106, Application US/09684855
; Patent NO. 6599735
; GENERAL INFORMATION:
; APPLICANT: P. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 13073
US-09-684-855-106

Alignment Scores:
Pred. No.:      0.00207      Length:      440
Score:          119.00      Matches:    102
Percent Similarity: 36.67%      Conservative: 70
Best Local Similarity: 21.75%      Mismatches: 183
Query Match:      5.07%      Indels:    114
DB:              4          Gaps:       24

US-10-034-985-1 (1-1323) x US-09-684-855-106 (1-440)
QY      i9  CCATTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTCGATTGCTCAGAGTGACCG 78
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```
Db      27  ProPhePheSerLeuGluAspGluLeuSerValSerSerLysLeuProLys----- 43
QY      79  GAGCTCAAGCTGAAAGTGTGGTGAATTGTCAGTCGTCATGTCGTGGTGGTGTCTCAAC- 135
          :::::  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      44  AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer 63
QY      136  -----AAGCCCAACGCAACTGATGTGAGGATGTCACCCACGACGCA----- 174
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      64  SerLysSerLysLysTyrLysLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 83
QY      175  -----TGGCCAACCTGGCCCGGTAAACCTGGGTGG-----CTGACA 210
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      84  LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 103
QY      211  CCGCGNGGTGGTGCAGTAATCGCTATCTCGA-----CATTACCAACGCGACGCTG 264
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      104  ProPheGlyGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla 123
QY      265  GTAGCCGAGCGGATTGCTGCGCAAAAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCT 324
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      124  LeuAlaArgSerValValPro-----PheIleArg 133
QY      325  GCTGATGTGACGAGGTACCCGTAAACAGGCGAAGCCTTCCGCCCGCGGCTG----- 378
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      134  AlaSerGlySerAspArgValIleAlaSerGlyGluLysPheIleGluGlyPheGln 153
QY      379  -----GCACCTGACTGTGCAATAACGTA 402
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      154  AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIleIle 173
QY      403  CATACCAGGCGAGATACGTCACGTCCTGATCCGCTTATTTAACTCTCTAAACCTGCGTT 462
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      174  ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 187
QY      463  TGCACAACTGGATAACCGAACGACGTGACGCGCATCTCAGAGGCGAGGAGGTCAATT 522
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      188  Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 200
QY      523  ---GCTGACTTTTACCGGGCATCGGCAACCGCGCTTTCGGAACCTGGAACGGGTCTAAT 579
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      201  AlaAlaAsnPheThr-----AlaLeuPhe 208
QY      580  TTTCCCGCAATCAAACTTGTGCTTAAACGTCAGAAACAGGACGAAAGCTGTTCTATTAACG 639
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      209  AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 226
QY      640  CAGCGATATTACCTCGGAACCTC-----AAGTGAGCGCGCGACAAATGCTCTCATTAACCGGT 693
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      227  AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaArgThrSer 246
QY      694  GCGGTAAGCTCGCATCAATGCTGACGGAGATATTTCTCTCGCAACAGCACAGGGAATG 753
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      247  AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 262
QY      754  CCGGAGCGCGGTGGGGAAGGATCACCAGATTTCACACGAGTGAACACCTTTGCTAAGTTTG 813
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      263  -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 272
QY      814  CATAACGCGCAATTTTATTGTTGTTACAAACGACCGCAGAGGTGCGCGCAGCGCGCCACC 873
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      273  -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 290
QY      874  CCGTTATTGATTGATCATGCGAGCGTTGACGCCCGCCCATCCACCGCAAAACAG----- 927
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      291  GlyPheThrAsnGlnLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 310
QY      928  -----CGGTATGTTGATGACATTACCCACTTCCAGTACTGTTTATTGGC 969
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      311  ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 330
QY      970  -----GGACACGATATACTCGGCAATCTCGCGCGCGCATCGGAGCTCAACTGGACG 1023
          ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Db      331  AspPheSerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 345
```


4

```

; TYPE: PRT
; ORGANISM: Aspergillus fumigatus 13073
US-09-488-265B-6

Alignment Scores:
Pred. No.: 0.00207 Length: 440
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 4 Gaps: 24

US-10-034-985-1 (1-1323) x US-09-488-265B-6 (1-440)
Qy 19 CCATTTTATCTCTTCTGATTCCGTAACCCCGCAATCTGCATTCGTCAGAGTGAGCCG 78
Db 27 ProphePheSerLeuGluAspGluLeuSerValSerSerLeuProLys----- 43
Qy 79 GAGCTGAAGCTGGAAGTGGTGAATGTCAGTCGTCATGTCATGTCGTCGTCGTCGTCGTCG 135
Db 44 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaAArgTyrProThrSer 63
Qy 136 -----AAGGCCACCACTGATGTCAGGATGTCACCCAGACGCA----- 174
Db 64 SerLysSerLysTyrLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 83
Qy 175 -----TGCCCAACCTGGCGCGGTAAACTGGGTTGG-----CTGACA 210
Db 84 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspLeuThr 103
Qy 211 CCGGNGGTGTGAGTAATCGCTATCTCGGA-----CAATACCAACGCGAGCGTGTG 264
Db 104 ProPheGlyGluGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla 123
Qy 265 GTACCCGACGATGTCGGCGAAGAGGCGTCCCGCAGTCGTGTCAGGTCGCGATTATT 324
Db 124 LeuAlaArgSerValPro-----PheIleArg 133
Qy 325 GCTGATGTCGACGAGCTACCGTAAACAGCGCAAGCTTCGCGCGGCGCTG----- 378
Db 134 AlaSerGlySerAspArgValIleAlaSerGlyLysPheIleGlyLysPheGlnGln 153
Qy 379 -----GCACCTGACTGTGCAATTAACCGTA 402
Db 154 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 173
Qy 403 CATACCCAGGAGATAGTCCAGTCCCGATCCGTTTAACTCTCTAAACCTGCGGTT 462
Db 174 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 187
Qy 463 TGCCAACTGGGATAACGGGAAGTGAAGTACGCGATCCTCAGCAGGCGAGGAGGTCATTT 522
Db 188 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 200
Qy 523 ---GCTGACTTTACGGGATCGGCAACAGCGGTTTCGCGAAGTGGGAGCGGTCCTTAAT 579
Db 201 AlaAlaAsnPheThr-----AlaLeuPhe 208
Qy 580 TTTCGGCAATCAACTGTGCTTAAAGCTGAGAAACAGGACGAAAGCTGTTCAATTAACG 639
Db 209 AlaPro-----AspIleArgAlaAlaGlnLysHisLeuProGlyValThrLeuThr 226
Qy 640 CAGGCATTACCATCGGAATCTC-----AAGGTGAGCGCCGACAAATGTCATTAAACCGGT 693
Db 227 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaAArgThrSer 246
Qy 694 GCGGTAAAGCTCGCATCAATGCTGACGAGATATTTCTCTGCAACAGCACAGGGAATG 753
Db 247 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 262
Qy 754 CCGGAGCGGGGTGGGAAGGATCACCAGATTACACAGTGGGAACACCTGCTTAAGTTTG 813
Db 263 -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 272
```

```

Qy 814 CATAACCGCGCAATTTTATTGTTTACACGACGCCAGAGGTGCCCCGACGCGCGCCACC 873
Db 273 -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 290
Qy 874 CCGTTATTGATTTGATCGGAGCGTTCAGCCCGCCCATCCACCGGCAAAACAG----- 927
Db 291 GlyPheThrAsnGluLeuIleAlaAArgLeuThrArgSerProValGlnAspHisThrSer 310
Qy 928 -----CGGTATGGTGTGACATTAACCCACTTCAGTACTGTTTATTATGCC 969
Db 311 ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 330
Qy 970 -----GGACACGATACTAATCGGCAATCTCGGCGCGCAGCTGGAGCTCAACTGGACG 1023
Db 331 AspPheSerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 345
Qy 1024 CTTCGCCGTGAGCGGATACACGCGCGCAGCGGTGGTGAACCTGGTGTGTAACGCTGGCGT 1083
Db 346 -----GlyLeuTyrAsnGlyThrGluProLeuSerArgThrSerValGluSerAlaLys 363
Qy 1084 CGGCTAAGCGAT---AACAGCCAGTGGATT-----CAGGTTTCGCTGCTTCCAG 1131
Db 364 GluLeuAspGlyTyrSerAlaSerTrpValProPheGlyAlaAArgAlaTyrPheGlu 383
Qy 1132 ACTTTACAGCAGATGCGTGATAAAACCGCGCTGTCAATTAATAACGCCCGCGGAGAGGTG 1191
Db 384 ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 402
Qy 1192 AAAGTACCTGCGCAGATGTAAGAGCGAATACGCGAGGCGCATGTTCTGTTGGCAGGT 1251
Db 403 ValValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 421
Qy 1252 TTTACGCAATCTGTAATGAAGCACGC 1278
Db 422 PheValLysGlyLeuSerTrpAlaArg 430
```

```

RESULT 15
US-09-044-718-12
; Sequence 12, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-12
```

```

Alignment Scores:
Pred. No.: 0.00209 Length: 449
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 4 Gaps: 24
```

US-10-034-985-1 (1-1323) x US-09-044-718-12 (1-449)

Qy 19 CCATTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTGCATTCGTCAGAGTGAGCCG 78

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:08:45 ; Search time 208.5 Seconds

(without alignments)
4908.240 Million cell updates/sec

Title: US-10-034-985-1

Perfect score: 2349

Sequence: 1 atgaagcgtatcttaatcccc.....atcaccatcaccatcaactaa 1323

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-O=/cgn2.1/USPTO.spool.p/US10034985/runat.10032005.164239.11943/app.query.fasta_1.1479
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985 @CGN 1.1 224 @runat.10032005.164239.11943 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	98.0	440	4 AAB37892	Aab37892 Escherich
2	2302	98.0	440	5 AAE22836	Aae22836 Escherich
3	2302	98.0	440	5 AAE15806	Aae15806 Escherich
4	2302	98.0	440	6 ADA19446	Ada19446 E. coli B
5	2302	98.0	440	8 ADO50292	Ado50292 Escherich
6	2282	97.1	440	8 ADO50304	Ado50304 Escherich
7	2279	97.0	440	8 ADO50302	Ado50302 Kangaroo
8	2235	95.1	432	3 AAB36257	Aab36257 Lama2/APP
9	2235	95.1	432	3 AAB36261	Aab36261 R15/APPA
10	2235	95.1	432	3 AAB36262	Aab36262 SV40/APPA

11	2235	95.1	432	3 AAB36259	Aab36259 R15/APPA
12	2235	95.1	432	3 AAB36258	Aab36258 R15/APPA
13	2235	95.1	432	3 AAB36263	Aab36263 Lama2/APP
14	2235	95.1	432	3 AAB36260	Aab36260 R15/APPA
15	2235	95.1	432	3 AAU77775	Aau77775 Phytase a
16	2235	95.1	432	5 AAE15807	Aae15807 Escherich
17	2235	95.1	432	5 ABP53625	Abp53625 Acid phos
18	2235	95.1	432	6 ADA19450	Ada19450 E. coli K
19	2235	95.1	432	6 ABU15427	Abu15427 Protein e
20	2235	95.1	432	7 ADC87745	Adc87745 Escherich
21	2235	95.1	432	8 ADL16129	Adl16129 Escherich
22	2235	95.1	432	8 ADO50298	Ado50298 Escherich
23	2230	94.9	432	4 AAE02631	Aae02631 E. coli a
24	2230	94.9	432	8 ADL16131	Adl16131 Shigella
25	2224	94.7	432	4 AAE02634	Aae02634 E. coli a
26	2224	94.7	432	8 ADL16130	Adl16130 Shigella
27	2223	94.6	432	4 AAE02635	Aae02635 E. coli a
28	2221	94.6	432	3 AAY94753	Aay94753 E. coli a
29	2221	94.6	432	6 ABR42161	Abr42161 Escherich
30	2210	94.1	432	4 AAE02632	Aae02632 E. coli a
31	2198	93.6	432	7 ADL02198	Adl02198 Appa. 5/2
32	2197	93.5	432	6 ABR42162	Abr42162 Escherich
33	2194.5	93.4	423	2 AAY01513	Aay01513 An Escher
34	2188	93.1	436	7 ADC87743	Adc87743 Modified
35	2171	92.4	432	8 ADO50300	Ado50300 Escherich
36	2141	91.1	430	5 AAE15808	Aae15808 Escherich
37	2141	91.1	430	6 ADA19451	Ada19451 E. coli K
38	2137.5	91.0	440	8 ADL16120	Adl16120 E. coli A
39	2131.5	90.7	440	8 ADL16125	Adl16125 Modified
40	2129.5	90.7	440	8 ADL16126	Adl16126 Modified
41	2129.5	90.7	440	8 ADL16122	Adl16122 Modified
42	2125.5	90.5	440	8 ADL16124	Adl16124 Modified
43	2121.5	90.3	440	8 ADL16128	Adl16128 Modified
44	2120.5	90.3	440	8 ADL16127	Adl16127 Modified
45	2112.5	89.9	440	8 ADL16123	Adl16123 Modified

ALIGNMENTS

RESULT 1

AAB37892

ID AAB37892 standard; protein; 440 AA.

XX AC AAB37892;

XX DT 07-MAR-2001 (first entry)

XX DE Escherichia coli B phytase enzyme.

XX KW Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

XX KW nutrition.

XX OS Escherichia coli.

XX PN WO2000071728-A1.

XX PD 30-NOV-2000.

XX PF 25-MAY-2000; 2000WO-US014846.

XX PR 25-MAY-1999; 99US-00318528.

XX PA (DIVE-) DIVERSA CORP.

XX PI Short JM, Kretz KA;

XX DR WPI; 2001-112081/12.

XX DR N-PSDB; AAC88885.

XX PT Improving the nutritional value of phytate-containing foodstuffs, using

XX PT phytase enzymes which catalyze the liberation of inorganic phosphate from

XX PT the phytates.

Claim 1; Fig 1; 147pp; English.

The present sequence is a phytase enzyme from *Escherichia coli* B. The enzyme catalyses the liberation of inorganic phosphate from the phytate in phytate-containing foodstuffs and can thus be used to improve the nutritional value of phytate rich ingredients

Sequence 440 AA;

Alignment Scores:

Pred. No.:	8,77e-230	Length:	440
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.00%	Indels:	0
DB:	4	Gaps:	0

US-10-034-985-1 (1-1323) x AAB37892 (1-440)

QY	1	ATGAAAGCGATCTTAAATCCCATTTTATCTCTCTTGATTCGGTTAAACCCCGCAATCTGCA	60
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGCTGGTGGATTGTCAGTCGTCATGGT	120
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleLeuSerArgHisGly	40
QY	121	GTGCGTGTCTCCAAACCAAGGCCACGCAACTGTATGCAGAGATGTCAACCCAGACGCATGGCCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGGCCGGTAAACTGGGTGGTGACACCGCGNGTGTGTGAGCTAATCGCTCTATCTC	240
DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIleLeu	80
QY	241	GGACATTACCAAGCCAGCGCTCTGTAGCCGACGAGTTCCTGGCGAATAAAGGGCTGCCCG	300
DB	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCTGGTCAGTCCGATATTGCTGATGTCGACGAGCGGTACCCGTAAACAGGCGAA	360
DB	101	GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCCCGCGGCTGGCACTGTGCAATAACCGTACATACCAGGCGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TTCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGGTTGCCCACTGGATAACGCG	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AAGTGCACGTAGCGCATCCTCAGCAGGCGCAGAGGGTCAATTGCTGACTTTTACCGGGCAT	540
DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
QY	541	CGGCAACCGCGTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC	600
DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGCGAGGCATTACCATCGGAATC	660
DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AAGTGAGCGCCGACATGTCTCATTAACCGGTGGGGTAAGCCTCGCATCAATGCTGACG	720
DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACACAGAGGAATGCGGAGCCGGGTGGGGAAGCATCAAC	780
DB	241	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
QY	781	GATTTCACACAGTGGAAACACCTTGCTAAGTTTGCATAACGGCGCAATTTTATTTGCTCAAA	840

RESULT 2

AAE22836
ID AAE22836 standard: protein: 440 AA.

AAE22836;

DT 21-AUG-2002 (first entry)

DE Escherichia coli phytase protein.

AA Dietary aid; biocompatible composition; therapeutic; digestive tract;
 KW foodstuff; digestion; phytase; enzyme.
 KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
 KW foodstuff; digestion; phytase; enzyme.

OS *Escherichia coli*.

PN WO200189317-A2.

29-NOV-2001

15-MAY-2001: 2001WO-US015764.

25-MAY-2000: 2000US-00580937.

PA (DIVE-) DIVERSA CORP.

XX
PI
Short JM.
Kretz KA.
O'donoghue E:

WPT: 2002-164149/21

DK N-FSDB; RAD364/3.
XX

PT comprise agent that assists in digestion, useful for delivering enzymes, PT therapeutics, medicine or agents to an organism.

PS Claim 5; Fig 1; 89pp; English.

The present invention relates to novel dietary aids comprising sustained release biocompatible composition which comprises an agent (enzymes such as phytase, amylase, esterase, protease) that assists in digestion. The

CC biocompatible composition is effective upon oral consumption and release
 CC in the digestive tract of a subject. The dietary aids are useful for
 CC delivering enzymes, therapeutics, medicine and agents to an organism. The
 CC use of enzymes and other agents in digestive aids of livestock or
 CC domesticated animals not only improves the animal's health and life
 CC expectancy but also assists in increasing the health of livestock or in
 CC the production of foodstuffs from livestock. The present sequence is
 CC Escherichia coli phytase protein
 XX
 SQ Sequence 440 AA;

Alignment Scores:

Prod. No.: 8,77e-230 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 5 Gaps: 0

US-10-034-985-1 (1-1323) x AAE22836 (1-440)

QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGTGTGATTGTCACTGCTCATGGT 120
 DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGGGTCTCAACCGAGCGCCAGCTGATCATGATGATGATGATGATGATGATGATGATGAT 180
 DB 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCGCGGTAAACATGGGCTTGGCTGACACCGCGGCTGCTGAGTAAATCGGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 QY 241 GGAACATTACCAACCGCAGCTCTGGTAGCCGAGGATTTGTCGCGAAAAAAGGGCTGCCCG 300
 DB 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGCTGGTCCAGTCGCGATTATGCTGATGTCGAGCGGTACCGGTAAACAGCGGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATACCCGTACATACCCAGGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGGCTTTGCCAACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGACTACCGGATCTCAGCAGGCGAGGAGGTCAATTCCTGACTTTACCGGCGAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGCAACAGCGGCTTCCGAACTGGACGGTGTCTTAATTTCCGCAATCAAACTGTGTC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGGCAATTACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAGCCCGACAAATGCTCTAATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATGCCGAGCGCGGGTGGGAAGGATCAAC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260

QY 781 GATTACACAGTGGGAACACCTTGCTAAGTTTCATATAACGCGCAATTTTATTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGCACCCAGAGGTTCGCCGACCGCGCCGCGCCGCGCCGCGCGCGCGCGCGCGCGCGCGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACCCCGCATCCACCGCAAAACAGGCGTATGGTGTGACATACCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGCGCACCATATACTTGGCAAAATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACCTTTCGCGGTGAGCGCGGATACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTAAGCGAATAACAGCCAGTTCGATTCAGTTTCGCTGCTTCCTTCAGACTTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAACGCGCGCTGCTCATTAATACCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGCGATGTTTCGTTGGCAGGTTTACGCAA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAACAGCACGATACCGCGCTGCAGTTTCAGATCTCATCCATCCCATCAC 1320
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis 440

RESULT 3
 AAE15806
 ID AAE15806 standard; protein; 440 AA.
 XX
 AC AAE15806;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Escherichia coli B phytase protein.
 XX
 DE Bacterial phytase; K12 appA phytase; protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; biopulping;
 KW non-alcoholic drink; biobleaching; B phytase.
 XX
 OS Escherichia coli.
 XX
 XX WO200109333-A2.
 PN
 XX
 XX 29-NOV-2001.
 XX
 XX 24-MAY-2001; 2001WO-US017118.
 PF
 XX
 XX 25-MAY-2000; 2000US-00580515.
 PR
 XX
 XX (DIVE-) DIVERSA CORP.
 PA
 XX
 XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue B;
 PI
 XX
 XX WPI; 2002-083108/11.
 DR
 XX
 XX N-PSDB; AAD25460.
 DR
 XX
 XX New bacterial phytase for e.g. improving the nutritional value of phytate
 PT -containing foodstuffs and subsequently improving the growth performance
 PT of an organism that consumes it, or in treating animal digestive systems.
 XX
 PS Claim 1; Fig 1; 170pp; English.

PR	13-APR-1999;	99US-00291931.	QY	121	GTCTGCTCCAAAGCCAGCACTGATGTCAGATGTCACCCAGCATGGCCA	180
PR	25-MAY-1999;	99US-00318528.	DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
XX	25-MAY-2000;	2000US-00580515.	QY	181	ACCTGGCCGTAATAACTGGTGTGGCTGACACCCGNGGTGGTGGAGCTAAATCGCTATCTC	240
PA	(SHOR/) SHORT J M.		DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuLeaLysLeu	80
PA	(KRET/) KRETZ K A.		QY	241	GGACATTACCAACGCCAGCGTCTGGTAGCGGATTCCTGGCGAAAAAGGGCTGCCCG	300
PA	(BART/) BARTON N R.		DB	81	GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
PA	(GARR/) GARRETT J B.		QY	301	CAGTCTGGTCCAGTCCGGATTATTGCTGATGTCGACGAGCGTACCCGTAACAGCGGAA	360
PA	(DONO/) O' DONOGHUE E.		DB	101	GlnSerGlyGlnValAlaLeileAlaAspValAspGluArgThrArgLysThrGlyGlu	120
XX	(MATH/) MATHUR E J.		QY	361	GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG	420
PI	Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;		DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
PI	O' Donoghue E, Mathur EJ;		QY	421	TCAGTCCCGATCCGTTATTATCTCTAAACCTGGCGTTCGCCAATCGGATAACGCG	480
XX	WPI; 2003-040002/03.		DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
XX	Isolated Escherichia coli polynucleotide encoding a modified phytase		QY	481	AACGTGACTGACCGGATCTCTACGACGCGCAGGAGGTCAATTCCTGACTTTACCGGCAT	540
PT	enzyme, useful in the production of animal feed, for improving the		DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180
PT	nutritional value of phytate-containing foodstuff and for enhancing		QY	541	CGSCAAACGCGGTTTCGCGAATCGGAACGGGTCTTAATTTTCGCAATCAAACTCTGCG	600
PT	digestion in humans and animals.		DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
XX	Claim 1; Fig 1; 62pp; English.		QY	601	CTTAAACGTCGAGAAACAGGACGAAAGCTGTTCATTAAACGACGACATTCACCATCGAACTC	660
CC	The invention relates to an isolated Escherichia coli polynucleotide		DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
CC	encoding a phytase appearing as ADA19450 and having amino acids		QY	661	AAGTGAGCCGCCACGATGCTCTATTAAACCGTGGGTAGCTCGCATCAATGCTGACG	720
CC	modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.		DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
CC	Alao included the E. coli appA gene ADA19449 (or an oligonucleotide		QY	721	GAGATATTTCTCTGCAACAGCACAGGAAATCCGAGCGCGGGTGGGAGGATCACCC	780
CC	derived from it) or its mutant sequence ADA19452, expression vectors,		DB	241	GluilePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
CC	host cells, a method of improving nutritional value of a phytate-		QY	781	GATTTCACACAGTGGAAACACTTGTCTGTAAGTTTCATCAACCGCAATTTATTTGCTACAA	840
CC	containing foodstuff by contacting the phytate-containing foodstuff with		DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln	280
CC	a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes		QY	841	CGACGCCAGAGGTGGCCCGCCAGCGCGGCCACCCCGTTATGATTTGATCATGGCAGCG	900
CC	the liberation of inorganic phosphate from the phytate in the phytate-		DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
CC	containing foodstuff), a method to produce an animal feed containing a		QY	901	TTGACGCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG	960
CC	microbial phytase (comprising culturing the plant cell, plant part or		DB	301	LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu	320
CC	plant under conditions where the nucleotide sequence is expressed and		QY	961	TTTATTGCGCGACACGATATAATCTGGCAAAATCTCGCGCGCGCACCTGGAGCTCAACTGG	1020
CC	converting the plant cells, plants or plant into a composition for animal		DB	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
CC	feed), a feed composition for animals (comprising the plant seeds, plant		QY	1021	ACGTTCCCGGTGAGCCGGATAAACGCCCGCAGGTGGTGAACCTGGTTTGAACGCTGG	1080
CC	cells, plant parts or plants of a transgenic plant which is modified to		DB	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
CC	contain an expression system which expresses a nucleotide sequence		QY	1081	CGTGGGTAGCGGATACAGCCAGTGGATTTCAGTTTCGTTGCTTCCAGACTTTACAG	1140
CC	encoding a phytase enzyme, a transgenic non-human organism whose genome		DB	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
CC	comprising a heterologous nucleic acid sequence encoding a polypeptide		QY	1141	CAGATGGGTGATAAAACCGCGCTGTCTATTAAATACGCCCGCGGAGAGGTGAACTGACC	1200
CC	having phytase activity. The phytase enzyme is useful for improving the		DB	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
CC	nutritional value of phytate-containing foodstuff, in the production of		QY	1201	CTGCGAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTCTGTTGGCAGGTTTTACGCA	1260
CC	animal feed and for enhancing digestion in humans and animals. The		DB			
CC	invented method improves thermal tolerance and protease stability. It		QY			
CC	also improves the feeding value of phytate rich ingredients. The present		DB			
CC	sequence represents E. coli B wild-type phytase.		QY			
XX	Sequence 440 AA;		QY			
SQ	Alignment Scores:		QY			
	Pred. No.:	8.77e-230	Length:	440		
	Score:	2302.00	Matches:	440		
	Percent Similarity:	100.00%	Conservative:	0		
	Best Local Similarity:	100.00%	Mismatches:	0		
	Query Match:	98.00%	Indels:	0		
	DB:	6	Gaps:	0		
	US-10-034-985-1 (1-1323) x ADA19446 (1-440)					
QY	1	ATGAAGCGATCTTAATCCCAATTTTATCTCTGATTCGGTTAACCCCGCAATCTGCA	60			
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20			
QY	61	TTCCCTCAGATGAGCCGAGCTGAAGCTGAAAGTGTGGTATGTGTCAGTCGTCATGGT	120			
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40			

Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCACCACCATCCGGCTGCGAGTTTGAGATCTCATCACCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 RESULT 5
 ADO50292
 ID ADO50292 standard; protein; 440 AA.
 XX AC ADO50292;
 XX DT 29-JUL-2004 (first entry)
 XX DE Escherichia coli B phytase enzyme.
 XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; enzyme.
 XX OS Escherichia coli; B.
 XX FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 XX PN US2004091968-A1.
 XX PD 13-MAY-2004.
 XX PF 20-JUN-2003; 2003US-00601319.
 XX PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX (SHORT) SHORT J M.
 PA (KRETT) KRETT K.
 PA (GRAY) GRAY K A.
 PA (BART) BARTON N R.
 PA (GARR) GARRETT J B.
 PA (ODON) O'DONOGHUE E.
 PA (MATH) MATHER E J.
 XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 PI Mather EJ;
 XX WPI: 2004-374952/35.
 DR N-PSDB; ADO50291.
 XX
 CC Producing phytase, involves providing nucleic acid derived from bacteria
 FT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 XX
 CC Claim 3; SEQ ID NO 2; 74pp; English.
 XX
 CC The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase

CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli B phytase enzyme.
 XX SQ Sequence 440 AA;
 Alignment Scores:
 Pred. No.: 8,77e-230 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 8 Gaps: 0
 US-10-034-985-1 (1-1323) x ADO50292 (1-440)
 QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTCAGCCGAGCTGAAGCTGGAAGTGTGGTGAATTTGTCAGTCTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGCGTGCTCCAAACCAAGCCACGCAACTGATCAGGATGTCAACCCAGCAGCATGCCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTppPro 60
 QY 181 ACCTGGCCGTAATAACTGGGTGGCTGACACCCGNGGTGGTGGAGTAATCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTrpLeu 80
 QY 241 GGACATTACCAAGCCAGCGCTCTGGTAGCCGAGATTGTCTGCGGAAAAGGGCTGCCCG 300
 Db 81 GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTCGCCGATTATTGTGATGTCCAGAGCGTACCCTGTAACAAACGCGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGCTGGCCACTGTGCAATAACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCCGATCCGTTATTATCTCTAAACCTGGGCTTTGCCCACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTGACCGATCTTCAGCAGGCGAGGAGGTCAATTCTGACTTTACCGGCGAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGTTTCGCGAAGTGGAAACGGGTGCTTAATTTCCGCAATCAAACTGTGCG 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerArgLeuLys 200
 QY 601 CTTAAACGTCAGAAACAGACGAAAGCTGTTCATTAAACGAGGAGCATTTACCATCGGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAGCGCCGACCAATGTCTCATTAACCGGTGGGTGAAGCCTCGCATCAATCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCCGAGCCGGGTGGGGAGGATCACC 780
 Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACACAGTGAACACCTTTGCTTAAGTTTCATAACGCGCAATTTTATTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
 QY 841 CGCACCCAGAGGTTCGCCGACCGCCGCCCGCTTATTGATTTGATCGGACGCG 900

Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACGCCCCATCACCGCAAAACAGCGGTATGGTGTGACATTACCACCTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACGACGATACATCTGCGCAAACTCGCGGCGCACTGGAGCTCACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGCTTCCCGGTGACGCGGATAACACGCGCCAGGTGGTGAACCTGGTGTGAACTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTGGCTAAGCGATACACGCCAGTGATTCAGTTTCGTGGTGTTCCTCCAGACTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTATTAACGCGCTGTCTATTAAATACGCGCGCGGAGAGGTCAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGAGGATGTGAAGAGCGAAATCGCAGCGCATGTGTTCTGGCAGGTGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTGAGATCTCATCACCATCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 6
 ADO50304
 ID ADO50304 standard; protein; 440 AA.
 AC ADO50304;
 XX
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Escherichia coli W phytase 875PH2 mutant enzyme.
 KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; mutant; mutin.
 XX Escherichia coli.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT Misc-difference 160 /note= "E. coli B phytase Ala replaced with Ser"
 FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"
 FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"
 FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"
 XX
 PN US2004091968-A1.
 XX
 PD 13-MAY-2004.
 XX
 XX 20-JUN-2003; 2003US-00601319.
 XX
 PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX
 XX (SHORT) SHORT J M.
 PA (KRETZ) KRETZ K.
 PA (GRAY) GRAY K A.

(BART/) BARTON N R.
 (GARR/) GARRETT J B.
 (ODON/) O'DONOGHUE E.
 (MATH/) MATHER E J.
 XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 PI Mather EJ;
 XX WPI; 2004-374952/35.
 DR N-PSDB; ADO50291, ADO50303.
 XX
 PT Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 XX
 PS Disclosure; Page: 74pp; English.
 XX
 CC The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli W phytase mutant enzyme. Note: The present
 CC sequence is not shown in the specification but has been derived from
 CC Escherichia coli B phytase ADO50292.
 XX
 SQ Sequence 440 AA;

Alignment Scores:
 Pred. No.: 1,06e-227 Length: 440
 Score: 2282.00 Matches: 436
 Percent Similarity: 99.32% Conservative: 1
 Best Local Similarity: 99.09% Mismatches: 3
 Query Match: 97.15% Indels: 0
 DB: 8 Gaps: 0

US-10-034-985-1 (1-1323) x ADO50304 (1-440)

QY 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGGTGATGTGTCAGTCGTATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTCGCTGCTCCAAACCAAGCCAGCAACTGATGATGATGATGATGATGATGATGATGAT 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGGTGGTGGTGGTGGTGGT 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
 QY 241 GGCATTACCAAGCCAGCGCTGCTGGTAGCGGAGGATGCTGCGGCAAAAGGGCTGCCG 300
 Db 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTCGCGATTTATTGCTGATGTCGAGCGGTACCCGCTAAACAGCGCAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGGCACCTGACTGTGTGCAATAACCGTATACCCAGCGGAGTACG 420

Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Qy 421 TCAGATCCCGATCGGTATTATTAATCTCTAAACAGTGGCGTTTGCACAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnSer 160
 Qy 481 AAGTGTACTACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaGlyPheThrGlyHis 180
 Qy 541 CGCAAAACGGCTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATACGAGCAGCATTTACCATCGGAATC 660
 Db 201 LeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Qy 661 AAGTGTAGCCCGACATGCTCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Qy 721 GAGATATTCTCTGCAACAGCACAGGGAATGCGGAGCCGGGTGGGGAAGATCAC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Qy 781 GATTACACAGTGGACACTTCTGTAAGTTTCATACGCGCAATTTTATTGTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 Qy 841 CGCACGCGCAGAGGTGCGCCGAGCGCGCCACCCCGTTATTGGATTGATCATGCGAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 Qy 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGATCATCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyxGlyValThrLeuProThrSerValLeu 320
 Qy 961 TTTATTGCGGACGACGATCTAATCTGGCAAACTCGCGCGCACCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Qy 1021 ACGTTCCTCCGTCAGCGGATAACACGCGCCGAGGTGGTGAACCTGGTTTGAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Qy 1081 CGTCGGCTAAGCATACACGCCAGTGGATTTCAGTTTCGTTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Qy 1141 CAGATGGTGTATAAACCGCGCTGTCTATTAAATACGCGCCGAGAGGTGAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Qy 1201 CTGCGCAGGATGTGAAGACGAAATTCGCGAGGATGTTGCTGTCGAGGATTTTACGCA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Qy 1261 ATGCTGAATAGCACGCATACCGGTGAGTTTGAGTCTCATCCATCCATCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 RESULT 7
 ID ADO50302 standard; protein; 440 AA.
 AC AC
 XX AC
 XX AC
 DT 29-JUL-2004 (first entry)
 XX Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.
 DE
 XX

KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; mutant; muten.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"
 FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"
 FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"
 FT Misc-difference 312 /note= "E. coli B phytase Gly replaced with Ser"
 XX US2004091968-A1.
 PN 13-MAY-2004.
 XX 20-JUN-2003; 2003US-00601319.
 XX 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX (SHORT) SHORT J M.
 PA (KRETZ) KRETZ K.
 PA (GRAY) GRAY K A.
 PA (BART) BARTON N R.
 PA (GARR) GARRETT J B.
 PA (ODON) O'DONOGHUE E.
 PA (MATH) MATHER E J.
 XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 PI Mather EJ;
 XX WPI; 2004-374952/35.
 DR N-PSDB; ADO50301.
 XX Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 PS Disclosure; Page; 74pp; English.
 XX The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Kangaroo rat Escherichia coli phytase mutant enzyme. Note:
 CC The present sequence is not shown in the specification but has been
 CC derived from Escherichia coli B phytase sequence ADO50292.
 XX Sequence 440 AA;
 SQ
 Alignment Scores: 2.17e-227 Length: 440
 Pred. No.:

Score:	2279.00	Matches:	436
Percent Similarity:	99.09%	Conservative:	0
Best Local Similarity:	99.09%	Mismatches:	4
Query Match:	97.02%	Indels:	0
DB:	8	Gaps:	0
US-10-034-985-1 (1-1323) x AD050302 (1-440)			
QY	1	ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA	60
DB	1	MetLysAlaIleuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla	20
QY	61	TTGGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGGTATGTCAGTCGTCATGCT	120
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40
QY	121	GTCCGTGCTCCAAACGAGGCCAGCAATGATCAGAGTGTACCCAGACGATGCGCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGCCCGGTAAACTGGGTTCGCTGCACACCGCGGTGGTGAAGTAAATCGCTATCTC	240
DB	61	ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
QY	241	GGACATTACCAACGCGAGCTGCTGAGCGGAGTTCGCGCAAAAGGGCTGCCCG	300
DB	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCGTGTCAGTCGCGATTATGTCGACGAGCGGTACCGGTAAACAGCGCA	360
DB	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCGCGCGGTGGCACCTGCTGTGCAATAACCGTACATACCCAGCGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCCAGTCCCGATCCGTTATTAATCTCTTAAACCTGGCGTTTGGCAACTGGATAACGG	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACGTGACTGACCGGTCTCAGCAGCGGAGGAGGTCAATTCGCTGACTTTACCGGCAT	540
DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaGlyPheThrGlyHis	180
QY	541	CGGCAACCGCGTTCGCGAATCGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC	600
DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGCATTTACATCGGAATC	660
DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AGGTGAGCGCCGACATGCTCATTAACGGTGGCGTAAGCCTCCGATCAATGCTGACG	720
DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACAGCACAGGAATCCCGAGCGCGGTGGGGAAGGATCAACC	780
DB	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
QY	781	GATTACACAGTGGAAACACTTGTCTAAGTTTGCATTAACCGCAATTTTATTTGCTACAA	840
DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
QY	841	CGACGCGCAGAGTTGCCCGACGCGGCCACCCGTTATGATTTGATTCATGGCAGCG	900
DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
QY	901	TTGACGCGCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTACCACCTTCAGTACTG	960
DB	301	LeuThrProHisProProGlnLysGlnAlaTyrSerValThrLeuProThrSerValLeu	320
QY	961	TTTATTGCCGACACGATCACTAATCTGGCAAAATCTCGCGCGCAGCTGGAGCTCAACTGG	1020

Db	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
QY	1021	AGCCTTCCCGGTCCAGCGGATAACCGCGCAGGTGGTGAACCTGGTGTGAAACGCTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
QY	1081	CGTCGGCTAAGCGATAACAGCCAGTGGATTACAGTTTCCTGCTTCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
QY	1141	CAGATCGGTGATAAACCGCGTGTCTATTAAATACGCCGCCGAGAGGTGAACCTGACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
QY	1201	CTGGCAGGATGTCAAGAGCGAAATGCGAGGCGCATGTTCGTTGGCAGGTTTACGCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
QY	1261	ATCGTGAATGAAGCAGCGCATACCGCGTGCAGTTTTCAGATCTCATCACCATCACCAC	1320
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis	440
RESULT 8			
AAB36257			
ID	AAB36257 standard; protein; 432 AA.		
XX	AAB36257;		
XX	(revised)		
DT	20-SEP-2003 (first entry)		
XX	Lama2/APPA plasmid translated sequence.		
DE	Transgenic animal; salivary protein; phytase; phosphorus; animal growth;		
KW	environmental pollution; pig.		
OS	Mus musculus.		
OS	Escherichia coli.		
OS	Chimeric.		
XX	WO200064247-A1.		
PN	02-NOV-2000.		
XX	20-APR-2000; 2000WO-CA000430.		
XX	23-APR-1999; 99US-0130508P.		
XX	(UYGU-) UNIV GUELPH.		
PI	Forsberg CW, Golovan S, Phillips JP;		
XX	WPI; 2000-687245/67.		
DR	N-PSDB; AAC68294.		
XX	Transgenic non-human animal for gastrointestinal tract specific		
PT	expression of a protein, preferably phytase, comprises a nucleic acid		
PT	sequence including a heterologous transgene construct encoding the		
PT	protein.		
XX	Disclosure; Fig 5; 152pp; English.		
PS	The present invention provides transgenic animals which produce desired		
CC	proteins, in this case pigs which expresses phytase in the salivary		
CC	gland. Low phytase production levels result in phytate in the diet being		
CC	excreted and causing phosphorus contamination in water, as well as		
CC	reducing the growth of animals. The invention provides a number of		
CC	transgenes containing the E. coli APPA phytase coding sequence. (Updated		
CC	on 12-SEP-2003 to standardise OS field)		
XX	Sequence 432 AA;		
SQ			

Alignment Scores:

Pred. No.: 8.11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36257 (1-432)

QY 1 ATGAACGATCTTAATCCATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 60
 DB 1 MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTGGCTCAGAGTGAGCCGAGCTGAAGCTGAAAGTGTGTGATTGTCAGTCTGATGTT 120
 DB 21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGGTGCTCCAAACCAAGGCCAGCAACTGATGTCAGGATGTCACCCAGACGATGGCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaIlePro 60
 QY 181 ACCTGGCCGTAATAACCTGGTGTGCTGACACCGCGNGTGTGAGCTAATCGCCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGTCGGCAAAAGGGCTGCCCG 300
 DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaTysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGGTCGGGATTTATGCTGATGTCAGCAGCGTACCCGTAATAACAGGCGAA 360
 DB 101 GlnSerGlyGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCGGTATTATTAATCTCTAAACCTGGCGTTCGCAACTGGATACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGACTCAGCGGATCCCTCAGCAGGCGAGGAGGTCAATGCTGACTTTTACCGGCGAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGGTTCGCAACCTGGAACCGGTGTTAATTTTCGCAATCAAACTGTGTC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATACGAGGATTTACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAAGCCGCAATGCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTGCAACAGCACAGGGAATCGCGAGCCGCGGTGGGGAAGATCAACC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 QY 781 GATTACACAGTGGAAACCTTGTCAAGTTTCATTAACCGCAATTTTATTTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 QY 841 CGCACGCGAGGTGTCGCGCAGCGCCACCCGCTTATTTGATTTGATCATGGCAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCGCCATCCACGCGAAAACAGGCGTATGGTGTGATTCACCATTCAGTACTG 960
 DB 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

QY 961 TTTATTCCGGACACGATATAATCTGGCAAAATCTCGGGCGCACTGGAGCTCAACTGG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTTCCCGGTGAGCCGGATACACCGCCGAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGCTAAGCAATACACCGCAGTTCAGTTTCAGTTTCGCTGCTTCAGACTTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAGCGGCTGCTCATTAATAATACGCCCGCGAGAGGTGAACCTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTGCTTGGCAGGTTTACGCCAA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCACGATACCGCGCTGCAGTTTG 1296
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 RESULT 9
 AAB36261
 ID AAB36261 standard; protein; 432 AA.
 AC AAB36261;
 DX XX
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 DX XX
 DE R15/APPA plasmid translated sequence.
 XX XX
 KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth; environmental pollution; pig.
 XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 PN WO200064247-A1.
 XX XX
 PD 02-NOV-2000.
 XX XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX XX
 PR 23-APR-1999; 99US-0130508P.
 XX XX
 PA (UYGU-) UNIV GUELPH.
 XX XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX XX
 DR WPI: 2000-687245/67.
 DR N-PSDB; AAC68298.
 XX XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX XX
 PS Disclosure; Fig 21; 152pp; English.
 CC CC
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

Db	301	LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerVal	320
Qy	961	TTTATTCCCGGACACGATCTAAATCTGGCAAAATCTCGCGCGCGACCTCGAGCTCAACTGG	1020
Db	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
Qy	1021	ACGTTTCCCGGTGAGCGCGATAACACGCCGCCAGGTGGTGAACCTGGTGTGTTGAACGCTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Qy	1081	CGTCGGCTAAGCCATAACACGCCAGTGATTCAGGTTTCGGTTCCTCCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Qy	1141	CAGATGGTGCATAAACCGCCGTGTCTATTAAATACGCCGCCGCGAGAGGTGAAACTGACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
Qy	1201	CTGCGAGGATGTGAAGACCGAAATGCGCAGCGCATGTTCGTTGGCAGGTTTTACGCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Qy	1261	ATCGTGAATCAAGCAGCATACCGCGGTGCAGTTTG	1296
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
RESULT 10			
AA	AB36262		
ID	AA36262	standard; protein; 432 AA.	
XX	AA36262;		
XX	XX		
DT	12-SEP-2003	(revised)	
DT	20-FEB-2001	(first entry)	
XX	SV40/APPA	plasmid translated sequence.	
DE	DE		
XX	XX		
KW	Transgenic animal;	salivary protein; phytase; phosphorus; animal growth;	
KW	environmental pollution;	pig.	
XX	XX		
OS	Simian virus 40.		
OS	Escherichia coli.		
OS	Chimeric.		
XX	XX		
PN	W0200064247-A1.		
XX	XX		
PD	02-NOV-2000.		
XX	XX		
PF	20-APR-2000;	2000WO-CA000430.	
XX	XX		
PR	23-APR-1999;	99US-0130508P.	
XX	XX		
PA	(UYGU-) UNIV GUELPH.		
XX	XX		
PI	Forsberg CW,	Golovan S,	Phillips JP;
XX	XX		
DR	WPI; 2000-687245/67.		
DR	N-PSDB; AAC68299.		
XX	XX		
PT	Transgenic non-human animal for gastrointestinal tract specific		
PT	expression of a protein, preferably phytase, comprises a nucleic acid		
PT	sequence including a heterologous transgene construct encoding the		
PT	protein.		
XX	XX		
PS	Disclosure; Fig 22;	152pp;	English.
XX	XX		
CC	The present invention provides transgenic animals which produce desired		
CC	proteins, in this case pigs which expresses phytase in the salivary		
CC	gland. Low phytase production levels result in phytate in the diet being		
CC	excreted and causing phosphorus contamination in water, as well as		
CC	reducing the growth of animals. The invention provides a number of		
CC	transgenes containing the E. coli APPA phytase coding sequence. (Updated		

SQ Sequence 432 AA;

Alignment Scores:
 Pred. No.: 8,11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36261 (1-432)

QY	1	ATGAAGCGCATCTTAATCCATTTTATCTCTCTGATTCGCTTAACCCCGCAATCTGCA	60
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTCCGCTCAGAGTCAGCCGAGCTGAAGCTGGAAGCTGGTATGTCAGTCGTCATGGT	120
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly	40
QY	121	GTCCGCTGCTCCAAACCAAGCCACGCAACTGATGCGAGGATGTCAACCCAGAGCATGGCCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACTGSCCGGTAAAACTGGTGGCTGACACCGCNGGTGGTGAAGTAAATCGCTATCTC	240
DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIleLeu	80
QY	241	GGCATTTACCAAGCCAGCGTCTGGTAGCGGAGGATGCTGGCGAAAAGGGCTGCCCG	300
DB	81	GlyHisIleThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro	100
QY	301	CAGTCCTGGTCAGTCGCCGATTATTGCTGATGTCGAGCGGTACCCGTAACCAAGCGCAA	360
DB	101	GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATACCGTATACCCAGCGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCCAGTCCCGATCCGTTATTATCTCTTAAACCTGGCGTTTGGCAACTGATAACCGC	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACGTGACTACCGCATCTCAGCAGCGCAGGAGGTCAATTGCTGACTTTACCGGGCAT	540
DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
QY	541	CGGCAACCGCGTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTGTGC	600
DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTAAACGTGAGAAAACAGACGAAGCTGTTATTAAACGAGGATTTACCATCGGAATC	660
DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AAGTGCAGCGCCACAAATGTCATTAAACCGGTGGTAAGCTCGCATCAATGCTGACG	720
DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACAAGCACAGGAATCCCGAGCCGGGTGGGGAGGATCAC	780
DB	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
QY	781	GATTTCACACAGTGGAAACCTTGTGAAGTTTGCATATAACCGCCAATTTTATGCTACAA	840
DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln	280
QY	841	CGCACCGCAGAGTTCGCCGACCGCCGCCACCCGTTATTCGATTTCATCATCGCAGCG	900
DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
QY	901	TTGACGCCCCCATCCACCGCAAAAACAGGGGTATGGTGTGACATTAACCACTTCAGTACTG	960

LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 TTTATTGCGCGACAGTACTAATCTGCAAAATCTCGCGCGCGACTCGAGCTCAACTGG 1020
 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 ACCTTCCCGTTCAGCCGAGTAAACCGCGCAGGTGTGAACCTGGTGTGTTGAACGCTGG 1080
 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 CGTCGGCTAAGCGATAACAGCCAGTGGATTGAGTTTCGCTGGTCTTCCAGACTTTACAG 1140
 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 CAGATCGGTGATAAACCGCGCTGTCTTAAATACCGCCCGCGAGAGGTGAAACTGACC 1200
 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 CTGCGCAGGATGTGAAGAGCGAAATCCGAGGCGCATGTCTTCTGGCAGGTTTACGCCAA 1260
 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1296
 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 10
 AAB36262
 ID AAB36262 standard; protein; 432 AA.
 AC AAB36262;
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE SV40/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 XX environmental pollution; pig.
 OS Simian virus 40.
 OS Escherichia coli.
 OS Chimeric.
 PN WO200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI; 2000-687245/67.
 DR N-PSDB; AAC68299.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX
 PS Disclosure; Fig 22; 152pp; English.
 CC
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated)

CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX
 SQ Sequence 432 AA;

Alignment Scores:

Pred. No.: 8.11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservatism: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36259 (1-432)

QY 1 ATGAACGATCTTAATCCATTTTATCTCTTCTGATTCGTTAAACCCCGCATCTGCA 60
 DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCCTCAGAGTACGCGAGCTGAAGTGAAGTGTGTGATGTGTCAGTGTGTCATGTT 120
 DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGGTCTCCAACACGCGCCACCACTGATGTCAGGATGTCACCCCGAGCGATGCGCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCGGTAAACCTGGTGTGCTGACACCGCGGTGTGAGTCAATCGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 QY 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAAGGGGTGCCCG 300
 DB 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTTCGGATATTGCTGATGTCGACGAGCGTACCCGTAAACAGGCGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGGACCTGACTGTGCAATACCGTACATACCCAGCGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTATTAATCTCTAAACCTGGCGTTTCCCACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTGACGATCTCTCAGAGCGGAGGAGGTCAATTGCTGACTTTACCGGCAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACGCGGTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTCGAGAACAGGAGAAAGCTGTTTCAATTAACGAGGATTAACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGGTGACGCGCAGATGCTCAATTAAACGGTGGGTAAAGCTTCGATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGTATTCTCTCGACACGACGAGGATGCGGCGCGGTGGGAGGATGATCACC 780
 DB 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACACGATGGAACCTTCTTAAGTTTGCATAACCGCGCAATTTTATTTCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlLeuLeuGln 280
 QY 841 CGCACCGCAGAGGTTGCCCGCAGCGCCACCCCGTTATTGGATTGATCGCAGCG 900

Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGGACACAGATACTAATCTGGCAAAATCTCGCGCGCGACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTCTCCCGGTGACGCGGATACACGCGCCAGGTGTCAACTGTGGTGTGAACTGGTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGCTAAGCGATAACACCCAGTGGATTTCAGTGTTCGCTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAGCGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGAGGATGTCAAGAGCGAAATGCGCAGGCGATGTTCTGTGGCAGGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATCAAGCAGCGATACCGCGGTGCGATTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 12
 AAB36258
 ID AAB36258 standard; protein; 432 AA.
 XX AAB36258;
 AC AAB36258;
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX WO200064247-A1.
 PN 02-NOV-2000.
 PD
 XX 20-APR-2000; 2000WO-CA000430.
 PF
 XX 23-APR-1999; 99US-0130508P.
 PR
 XX (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 DR N-PSDB; AAC68295.
 DR
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 PS Disclosure; Fig 18; 152pp; English.
 XX
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary

CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX
 SQ Sequence 432 AA;

Alignment Scores:
 Pred. No.: 8,11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservatives: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36263 (1-432)

QY 1 ATGAAAGCGATCTTAATCCCAATTTTATCTCTGATTCCTGTAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTCGTCTAGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTCCGTGCTCCAAACGAGCCAGCACTGATCAGGATGTACCCAGCAGCATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCCGTAATAACTGGGTGTGGCTGACACCGCGNGGTGGTGAAGTAACTCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGAATTAACCAACCGCAGCTCTGTTAGCCAGCGATTTGTCGGCAAAAGGGCTGCCCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGCTGGTCCAGTCCGATTTATTTCTGATGTGACGAGCGTACCCGTAACAGCGGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTTATCTCTAAACCTGGCGTTTCCCAACTGGATACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGCTGACTGACCGATCTCCACAGCGGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGATTAACCATCGGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGTGCGCCGACATGCTCATTAACCGGTGCGGTAGCCTCCGATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGGGGTGGGGAAGGATCAACC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTCACACCGAGTGAACACTTGTGTAAGTTTGTGATTAACCGCAATTTTATTTGCTACAA 840

Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGln 280
 QY 841 CGCACCCAGAGGTTCCCGCGCAGCCGCCACCCCGTTATTGATTTGATTCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGACCATATACTATCTGGCAAAATCTCGCGCGCAGCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACCTTTCCCGGTGAGCCGGATACACCGCCCGCAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTAAGCATTAACAGCCAGTTCAGGTTTCGCTGCTCTTCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAGCGCGCTCATTAATAACGCCGCCGAGAGGTGAACGAC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTGTTGTCGAGGTTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 14
 AAB36260
 ID AAB36260 standard; protein; 432 AA.
 XX
 AC AAB36260;
 XX
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX
 DE Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 KW
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX
 PN W0200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI; 2000-687245/67.
 DR N-PSDB; AAC68297.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX

XX SQ Sequence 432 AA;
Alignment Scores:
pred. No.: 8.11e-223 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservativeness: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAU77775 (1-432)

QY	1	ATGAAGCGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCTCGCAATCTGCA	60
Db	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTGCTCAGAGTCAGCCGGAGCTGAGCTGAAAGTGTGGTATGTCAGTCTGCTATGCT	120
Db	21	PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerArgHisGly	40
QY	121	GTGCTGCTCAACCAAGGCCAGCAACTGATGAGGATGTCACCCAGACGATGCGCCA	180
Db	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGCGGTAAACACTGGGTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC	240
Db	61	ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
QY	241	GGACATTACCAAGCCAGCTCTGTTAGCCAGCGATGCTGGCGAAGAAAGGCTCGCCG	300
Db	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCTGCTCAGGTCGCGATTATTGCTGATGTGCACGAGCGTACCCGTAACACAGGCGAA	360
Db	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG	420
Db	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCCAGTCCCGATCCGTTTATTAATCTTAACACTGGCGTTCGCAACTGGATACGCG	480
Db	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGGTCAATTGCTGATTTTACCGGCGAT	540
Db	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180
QY	541	CGCAAAACGCGGTTCCGCAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTC	600
Db	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTAACGTCAGAAACAGGAGAAAGCTGTTCATTACGAGGAGCATACCATCGGAATC	660
Db	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AAGTGAGCCCGCAATGCTCATTAAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG	720
Db	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTCTCTGCAACAGCAAGGAAATCCGAGCCGGGTGGGAAGGATCACC	780
Db	241	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr	260
QY	781	GATTACACAGTGGAAACCTTGCTAAGTTTGATTAACCGGCAATTTTATTTGCTACAA	840
Db	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
QY	841	CGCACGCCAGAGTTGCCCGCAGCCGCCACCCCGTTATTGGATTGATCATGGCAGCG	900
Db	281	ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300

Search completed: March 11, 2005, 23:28:07
Job time : 237.5 secs

QY	901	TTGACGCCCCATCCACGGCAAAACAGGCGTATGTTGATATCCCACTTCCAGTACTG	960
Db	301	LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
QY	961	TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCACTGGAGCTCAACTGG	1020
Db	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
QY	1021	ACGCTTCCCGGTGAGCCGATACACCGCCGAGGTGGTGAACCTGGTGTGTTGAAACGCTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
QY	1081	CGTCGCTAAGCGATACAGCCAGTTCAGGTTTCGCTGCTTCCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
QY	1141	CAGATCGGTGATAAAACGCGCTGTCTATTAAATACGCGCCGAGAGGTGAAACTGACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
QY	1201	CTGCGAGGATGTGAAGAGCAATGCGCAGGCGCATGTGTTCTGTTGCGAGGTTTACGCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
QY	1261	ATCGTCAATGAACACGATACCGCGCTGCGAGTTTG	1296
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432

THIS PAGE BLANK (USE REVERSE)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:17:45 ; Search time 55.5 Seconds
(without alignments)
4587.199 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatcc.....atcaccatcaccatcactaa 1323

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p,model -DEV=xlp
-Q=/cgr2_1/USFTO_spool_p/US10034985/runat_10032005.164240.11959/app_query.fasta_1.1479
-DB=PIR_79 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985 @CGN 1 71 @runat_10032005.164240.11959 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	95.1	432	2 B36733	acid phosphatase (
2	2210	94.1	434	2 H90770	phosphoanhydride p
3	2210	94.1	444	2 D85633	hypothetical prote
4	943.5	40.2	441	2 AC0201	acid phosphatase (
5	545	23.2	413	2 F90773	periplasmic glucos
6	545	23.2	413	2 D85636	periplasmic glucos
7	541	23.0	417	2 S25627	glucose-1-phosphat
8	539	22.9	413	2 JY0087	glucose-1-phosphat
9	526	22.4	414	2 E87316	periplasmic phosph
10	524	22.3	413	2 AG0632	glucose-1-phosphat
11	139.5	5.9	423	1 S06157	acid phosphatase (
12	133	5.7	416	2 T16058	hypothetical prote
13	131	5.6	423	2 A33395	acid phosphatase (
14	122.5	5.2	421	2 S14742	acid phosphatase (

15	117	5.0	386	1 JH0610	acid phosphatase (
16	116.5	5.0	450	2 A54429	paired box transcr
17	113.5	4.8	344	2 B89130	protein F52E1.8 [i
18	111	4.7	2554	1 TVFF7L	kinase-related pro
19	110	4.7	459	2 S52250	paired box transcr
20	109.5	4.7	1662	2 T18540	mofA protein precu
21	108.5	4.6	10223	2 T30225	polyketide synthas
22	107.5	4.6	457	2 A56925	paired box transcr
23	107	4.6	1048	1 BVEGSC	exonuclease (EC 3.
24	106.5	4.5	441	2 H70632	hypothetical prote
25	104.5	4.4	2535	2 AC0304	probable hemolysin
26	102	4.3	369	2 JQ2278	hydroxymethylbilan
27	102	4.3	638	2 A29440	signal recognition
28	102	4.3	1047	2 C85535	ATP-dependent dADN
29	102	4.3	1047	2 G90684	ATP-dependent dADN
30	101.5	4.3	397	2 C81716	hypothetical prote
31	101.5	4.3	479	1 JN0890	acid phosphatase (
32	101.5	4.3	524	2 H75530	probable acid-CoA
33	101.5	4.3	769	2 B87681	tyrosine kinase Di
34	101.5	4.3	4135	2 T42629	tenascin-X - bovin
35	101	4.3	360	2 JE0116	zinc-finger protei
36	101	4.3	26926	1 I38344	titin, cardiac mus
37	100.5	4.3	438	2 S64682	acid phosphatase (
38	100.5	4.3	478	2 S58315	major plasmid tran
39	100.5	4.3	1214	2 AG2897	conserved hypochet
40	100.5	4.3	1387	2 A97673	probable periplasm
41	100.5	4.3	6260	2 T30228	polyketide synthas
42	100	4.3	328	2 A83572	pyridoxal phosphat
43	100	4.3	1374	1 VCB517	major capsid prote
44	100	4.3	3828	2 T13857	trithorax protein
45	99	4.2	636	2 G81983	pillin glycosylatio

ALIGNMENTS

RESULT 1

B36733
N; Acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
C; Species: Escherichia coli
C; Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C; Accession: B36733; S18018; B64839; A26534; S17960; S33278
R; Basset, J.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 5497-5500, 1990
A; Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals eign
A; Reference number: A36733; MUID:90368616; PMID:2168385
A; Accession: B36733
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-432 <DAS>
R; Greiner, R.; Jany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A; Title: Characterization of a phytase from Escherichia coli.
A; Reference number: S18018
A; Accession: S18018
A; Molecule type: protein
A; Residues: 23-33 <GRE>
R; Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A64720; MUID:97426617; PMID:9278503
A; Accession: B64839
A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-432 <BLAT>
A; Cross-references: GB:AE000200; GB:U00096; NID:g2367111; PIDN:AAC74065.1; PID:g1787215;
A; Experimental source: strain K-12, substrain MG1655
R; Touati, E.; Danchin, A.
Biochimie 69, 215-221, 1987
A; Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phosph
A; Reference number: A26534; MUID:87271766; PMID:3038201

A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-50; 'NAGCHPRRMANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>
A;Cross-references: GB:X05471; NID:g40925; PIDN:CRA29031.1; PID:g40927
R;Passa, J.; Fsihi, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A;Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a putative
A;Reference number: S17958; MUID:92049231; PMID:1658595
A;Accession: S17960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-17 <DA2>
A;Cross-references: GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:g238659
R;Greiner, R.; Konietzny, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from Escherichia coli.
A;Reference number: S33278; MUID:93256556; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31, 'A', 33-34 <GR2>
C;Comment: In addition to CAMP-mediated control, this enzyme is induced when bacterial
naerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mon
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Alignment Scores:
Pred. No.: 8 87e-175 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x B36733 (1-432)

QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTGATCCGTTAACCCCGCAATCTGCA 60
DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTGCTCAGAGTACGGCGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTCCTCATGTT 120
DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGTCTCAACCAAGGCGACGCACTGATGCGAGATGTACCCCGACGCGATGCGCA 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
QY 181 ACCTGCGCGTAAACTGGTGTGGCTGTGACACCGCGGGTGGTGAAGTAAATCGCCTATCTC 240
DB 61 ThrTyrProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTACCAAGCGCGCTGTGTAGCCGCGAGATGTCGCGGAAAAGGGCTGCGCG 300
DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTGTGTGTCAGTTCGGATTATTGCTGATGTCGAGCGGTACCCGTTAAACACAGCGAA 360
DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGTGGACCTGATGTGCAATAACCGTATACCCAGGAGATACG 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCGTTATTATTAATCTCTAAACTGGCGTTTGGCCACTGGATAAGCG 480
DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

QY 481 AAGCTGACTGACGCGATCTCTCAGCGGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGGCTTTCGCAACTGGAACCGGTGCTTAATTTCCGCAATCAACTTTGTGC 600
DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTTGAGAACACGAGCAGCAAGCTGTTCATTAAACGAGGCATTACCATCGGAATC 660
DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTTCAGGCGCGACAACTCTCTATTAAACGCGTAAAGCTCGCATCAATGCTGACG 720
DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCAGCAGGAATGCGGAGCGCGGGTGGGGAAGGATCACC 780
DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTCAACAGTGGGAACACCTTGTAACTTTGTGATAACGCGCAATTTATTATTGCTACAA 840
DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCCAGAGGTTCGCCGAGCGCGCCACCCCGTTATTGATTTGATCATGCGAGCG 900
DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGGGGTATGGTGTACATTACCCACTTCAGTACTG 960
DB 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGCGACGATACTTAATCTGCAAAATCTCGCGCGCGCATCTGAGCTCAACTGG 1020
DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTTCCCGTTCAGCCGATTAACCGCGCGAGGTGGTGAACCTGTTGTGAACCTGG 1080
DB 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGCTAAGCGATTAACGCGAGTTCAGTTTCGTTCTCCAGCTTTACAG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACCGCGCTGTCTATTAAATACCGCGCGAGGTGAACCTGACC 1200
DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATGTAAGACGCAAAATGCGCAGGCGCATGTGTGCTGGCAGGTTTACGCAA 1260
DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACGACGATACCGCGCTGCGAGTTTG 1296
DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 2
H90770
phosphoanhydride phosphorylase [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; GB:BA000007; PIDN:BA834559.1; PID:g13360596; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:
A;Gene: ECs1136

Alignment Scores:

Pred. No.: 9.87e-173 Length: 434
Score: 2210.00 Matches: 427
Percent Similarity: 98.39% Conservatives: 0
Best Local Similarity: 98.39% Mismatches: 5
Query Match: 94.08% Indels: 2
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x H90770 (1-434)

```

Qy 1 ATGAACGATCTTAATCCATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAAGTGTGGTATTGTCAGTCGT 114
Db 21 PheAlaGlnSerGluProGluProGluLeuLeuLeuSerValValIleValSerArg 40
Qy 115 CATGGTGTGGCTGCTCCAAACCAAGCCACCACTGATGCAGGATGTCAACCCAGAGCA 174
Db 41 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 60
Qy 175 TGCCCAACCTTGGCCGGTAAACCTGGTGTGCTGACACCGCGNGGTGGTGAAGTAAATGCC 234
Db 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 80
Qy 235 TATCTCGGACATTACCAACGCCAGCGTCTGGTAGCCGACGAGTTCGGCGMAAAGGCC 294
Db 81 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 100
Qy 295 TGCCCGCAGTCTGTGAGTTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAACAA 354
Db 101 CysProGlnProGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThr 120
Qy 355 GGCBAACCTTGGCCCGCGGTGGACCTGACTGTGCAATAACCGTACATACCCAGGCA 414
Db 121 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 140
Qy 415 GATACGTCGATCCGATCCGTTATTTAATCCTTAATAACTGCGTTCGCCACTGGAT 474
Db 141 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 160
Qy 475 AACGCGAAGTGTGACTGACGGATCCTCAGCAGGCGAGGCTCAATTGCTGACTTTACC 534
Db 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 180
Qy 535 GGGCATCGCAACCGCGTTTCGCGAATCGGAACGGGTGCTTAATTTTCGCAATCAAC 594
Db 181 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 200
Qy 595 TTGTGCTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTAACGAGGATACCATCG 654
Db 201 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 220
Qy 655 GAATCAAGTGTAGCGCGCAATGTCTCAATTAACGGTGGGTGAAGCCTCGCATCAATG 714
Db 221 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
Qy 715 CTGACGGAGATATTTCCTCGTCAACCAAGCACAGGAAATGCCGAGCGCGGTGGGAAGG 774
Db 241 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 260
Qy 775 ATCAACGATTCACACGATGGAACACCTTGTCTGAATTTGCTAATACCGGCAATTTATTG 834
Db 261 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 280
Qy 835 CTACAACGACGCGCAGAGTTGCCCGCAGCGCCACCCGCTTATTGGATTTGATCATG 894
Db 281 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 300

```

```

Qy 895 GCAGCGTTGACGCCCATCCACCGCAAAACAGCGCTATGGTGTGACATTACCACTTCA 954
Db 301 IleAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 320
Qy 955 GTACTGTTTATTCGCGACACGATACTAATCTCGAAATCTCGCGCGCGACCTGGAGCTC 1014
Db 321 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 340
Qy 1015 AACTGACCGTTCCTCCGTCAGCGGATAAACACCCCGCAGGTGGTGAACCTGGTGTGAA 1074
Db 341 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 360
Qy 1075 CGCTGCGCTCGCTAAGCATACACCGCAGTCAGTTCAGTTCGCTGCTTCCAGACT 1134
Db 361 ArgTrpArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 380
Qy 1135 TTACAGCAGATGGTGATAAACCGCGCTGTCTAATAATACCGCGCGAGAGGTGAAA 1194
Db 381 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 400
Qy 1195 CTGACCTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTCGTTGGCAGGTTT 1254
Db 401 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 420
Qy 1255 ACCCAATCGTGAATGAGCAGCATACCGCGCTGCAGTTTG 1296
Db 421 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 434

```

RESULT 3

D85633
hypothetical protein appA [imported] - Escherichia coli (strain O157:H7, substrain EDL93:
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; GB:AE005174; MID:g12514245; PIDN:AAG55528.1; GSPDB:G1
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Alignment Scores:
Pred. No.: 9.9e-173 Length: 444
Score: 2210.00 Matches: 427
Percent Similarity: 98.39% Conservatives: 0
Best Local Similarity: 98.39% Mismatches: 5
Query Match: 94.08% Indels: 2
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x D85633 (1-444)

```

Qy 1 ATGAACGATCTTAATCCATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
Qy 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAAGTGTGGTATTGTCAGTCGT 114
Db 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 50
Qy 115 CATGGTGTGGCTGCTCCAAACCAAGCCACCGCACTGATGCAGGATGTCAACCCAGAGCA 174
Db 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
Qy 175 TGCCCAACCTTGGCCGGTAAACCTGGTGTGCTGACACCGCGNGGTGGTGAAGTAAATGCC 234
Db 71 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 90

```

```
QY 235 TATCTCGACATTACCAACGCCAGCGTCTGTAGCCGACGGATTGCTGGCGAAAGGGC 294
Db 91 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysGly 110
QY 295 TGCCCGCAGTCTGTCAGGTCGCGATTATTCTGATCTGCAGCAGCGTACCCGTAACA 354
Db 111 CysProGlnProGlyGlnValAlaIleAlaAspValAlaAspGluArgThrArgLysThr 130
QY 355 GCGAAGCCTTCGCGCGCGCTGCGACCTGCTGCTGCAATAACCGTACATACCAGGCA 414
Db 131 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 150
QY 415 GATACGTCAGTCCGATCCGTTATTTAACTCTTAACTGCGGTTTCGCAACTGGAT 474
Db 151 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 170
QY 475 AACCGAAGTGTACTGACGCGATCCTCAGCAGGCGAGGGTCAATTGCTGACTTTACC 534
Db 171 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 190
QY 535 GGGCATCGCAACCGGCTTCGGAACCTGCGAAGCGGTCTTAATTTTCGCAATCAAC 594
Db 191 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 210
QY 595 TTGTGCTTAAACGTGAGAAACAGCAGCAAGCTGTTCAATTAACGACGAGCATACCATCG 654
Db 211 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 230
QY 655 GAATCAAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGTAAACCTCGCATCAATG 714
Db 231 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 250
QY 715 CTGACGAGATATTTCTCTCCACACAGCAGGGAATCGGAGCGGGTGGGAGG 774
Db 251 LeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrpGlyArg 270
QY 775 ATCAACCATTCACACAGTGGAAACACCTGTGTAAGTTTGATACACGGCGCAATTTATTG 834
Db 271 IleThrAspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 290
QY 835 CTACAAACGACGCGCAGAGTTGCGCGAGCGCGCCACCCCGTTATTGATTTGATCATG 894
Db 291 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 310
QY 895 GCACGGTTGACGCCCATCCACCGCAAAACAGCGCATGTGTGACATTACCCACTTCA 954
Db 311 IleAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 330
QY 955 GTACTGTTTATTGCGGACACGATACTAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTC 1014
Db 331 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 350
QY 1015 AACTGAGCGCTTCCCGTACCGCGATTAACACGCGCGCAGTGTGTGACTGGTGTGAA 1074
Db 351 AsnThrThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 370
QY 1075 CGCTGGCGTGGCTAAGCGATAACAGCAGTGGATTTCAGGTTTCGCTGCTTCCAGACT 1134
Db 371 ArgThrArgArgLeuSerAspAsnSerGlnThrPheGlnValSerLeuValPheGlnThr 390
QY 1135 TTACAGCAGATGCTGTATAAACCGCGTGTCTATTAAATACGCGCGCGAGAGGTGAAA 1194
Db 391 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 410
QY 1195 CTGACCCCTGGCAGATGTGAAGACGAAATACGCGAGGCGCATGTGTTGTCGGCAGGTTT 1254
Db 411 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 430
QY 1255 ACGCAATCGTGAATGAGCAGCATACCGCGCGTCAAGTTTG 1296
Db 431 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 444
```

RESULT 4

AC0201
acid phosphatase (EC 3.1.3.2) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0201
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, F
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <R>
A;Cross-references: UNIPROT:Q8ZFP6; GB:AL590842; PIDN:CAC90470.1; PID:gl5979685; GSPDB:G
C;Genetics:
A;Gene: YPO1648
C;Keywords: phosphoric monoester hydrolase

Alignment Scores:
Pred. No.: 4,5e-69 Length: 441
Score: 943.50 Matches: 199
Percent Similarity: 60.78% Conservative: 66
Best Local Similarity: 45.64% Mismatches: 156
Query Match: 40.17% Indels: 15
DB: 2 Gaps: 7

US-10-034-985-1 (1-1323) x AC0201 (1-441)

```
QY 13 TTAATCCCAATTTTATCTCTCTGATTCCGTTAAACCCGCAATCTGCAATTCGTCAGAGT 72
Db 13 LeuValLeuMetLeuSerGlyLeuAlaAlaIleThrAlaProValAla-----Ala 29
QY 73 GAGCCGAG---CTGAAGCTGGAAGAGTGTGATGTCAGTGTCTAGTGTCTGTGTGCGTGT 129
Db 30 GluProSerGlyTyrThrLeuGluArgValIleLeuSerArgHisGlyValArgSer 49
QY 130 CCAACCAAGCCACGCAACTGATGCGAGGATGTCAACCCAGCATCGGCCAACCTGGCCG 189
Db 50 ProThrLysGlnThrGlnLeuMetAsnAspValThrProAspLysTyrProGlnTyrPro 69
QY 190 GTAAACTGGGTTGGCTGACACCGCGNGGTGGTGTGAGTAATACGCTATCTCGACATTC 249
Db 70 ValLysAlaGlyTyrLeuThrProArgGlyAlaGluLeuValThrLeuMetGlyGlyPhe 89
QY 250 CAACGCCAGCGTCTGTAGCCGAGGATTCCTGGCGAAAAGGGCTGCCGCGAGTCTGTGT 309
Db 90 TyrGlyAspTyrPheArgSerLeuGlyLeuAlaAlaGlyCysProAlaGluGly 108
QY 310 CAGTGTGCGATTATTGCTGATGTCGACGAGCGTACCGTAAACAGCGAGCGCTTCGCC 369
Db 109 GlyValTyrAlaGlnAlaAspIleAspGlnArgThrArgLeuThrGlyGlnAlaPheLeu 128
QY 370 GCCGGGTGGCACCTGCTGCAATAACCGTATACCATACCAGGAGAGATACGTCAGTCCC 429
Db 129 AspGlyValAlaProGlyCysGlyLeuThrValHisAsnGlnAlaAspLeuLysThr 148
QY 430 GATCCGTTATTATCTCTAAACCTGGGCTTGGCACTGGATACGCGAAGCGTACT 489
Db 149 AspProLeuPheHisProValGluAlaGlyValCysLysLeuAspAlaAlaGlnThrAsp 169
QY 490 GAGCGCATCTCAGCAGGCGAGGTCATATGCTGACTTTTACCGGCGCATCGGCAACG 549
Db 169 LysAlaIleGluGlnLeuGlyGlyProLeuAspThrValSerGlnArgTyrAlaLys 188
QY 550 GCGTTTCGCAATCGGAACGGTGTCTTAATTTCCGCAATCAAACTTGTCGCTTAAACGT 609
Db 189 ProPheAlaGlnMetGlyAspValLeuAsnPheAlaAlaSerProTyrCysLysSerLeu 208
QY 610 GAGAAACAGGACGNAAGCTGTTTCATTATACCGCAGGCATTACCATCGGAACCTCAAGTGCAGC 669
Db 610 GAGAAACAGGACGNAAGCTGTTTCATTATACCGCAGGCATTACCATCGGAACCTCAAGTGCAGC 669
```

Db 209 GlnGlnGlnGlyLysThrCysAspPheAlaHisPheAlaAlaAsnGluValAsnValAsn 228
Qy 670 GCCGAC-----AATGCTCATTAACCGGTCGCGTAAGCCTCGCATCAATGCTCAACGGAG 723
Db 229 LysGluGlyThrLysValThrLysSerGlyProLeuAlaLeuSerThrLeuGlyGlu 248
Qy 724 ATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGATCACCGAT 783
Db 249 IlePheLeuLeuGlnAsnAlaGlnAlaMetProGluValAlaTrpGlnArgLeuLysGly 268
Qy 784 TCACACAGTGAACACTGCTTAAGTTTGATACCGCAATTTTATTTGCTTACAACGC 843
Db 269 AlaGluAsnTrpValSerLeuLeuSerLeuHisAsnAlaGlnPheAsnLeuMetAlaLys 288
Qy 844 AGCCAGAGGTGCGCCGAGCGCGCCACCCGCTTATTGATTTGATCATGGCAGCGTTG 903
Db 289 ThrProTyrIleAlaArgHisLysGlyThrProLeuLeuGlnIleAspThrAlaLeu 308
Qy 904 ACGCCCCATCACCGCAAAACAGCGGTATGTTGATGACATTACCCTTCA-----954
Db 309 Thr-----LeuGlnLeuAspAlaGlnGlyGlnLysLeuProIleSerAlaGlnAsn 325
Qy 955 --GTACTGTTTATTGCGGACACGATACTAATCTGGCAATCTCGCGCGCACTCGAG 1011
Db 326 ArgValLeuPheLeuGlyGlyHisAspThrAsnIleAlaAsnIleAlaGlyMetLeuGly 345
Qy 1012 CTCAACTGGAGCTTCGCGTACCGGATAACACGCGCGAGTGTGTAAGTCTGTTT 1071
Db 346 AlaAspTrpGlnLeuProGluGlnProAspAsnThrProProGlyGlyGlyLeuValPhe 365
Qy 1072 GAACGCTGGCTCGGCTAAGCGATAACAGCAGTGGATTCAGTTTCGCTGCTTCCAG 1131
Db 366 GluLeuTrpGlnAsnProAspAsnHisGlnArgTyrValAlaValLysMetPheTyrGln 385
Qy 1132 ACTTACAGCAGATGCGTGATAAAACCGCGTGCATTA--ANTACGCCCGCGGAG 1188
Db 386 ThrMetAspGlnLeuArgAsnAlaGluLysLeuAspLeuLysAsnAsnProAlaGlyIle 405
Qy 1189 GTGAACCTGACCTCGGAGGATGTGAAGAGCGAAATCGCGGCGATGTTCGTGGCA 1248
Db 406 IleSerValAlaValAlaGlyCysGluAsnAsnGlyAspAspLysLeuCysGluLeuAsp 425
Qy 1249 GGTTTTACGCAATCGTGAATGAAGCAGCAGCATACCGCGTGCAGTTTG 1296
Db 426 ThrPheGlnLysLysValAlaLysValIleGluProAlaCysHisIle 441

RESULT 5
P90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: P90773
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; UID:21156231; PMID:11258796
A:Accession: P90773
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <HAY>
A:Cross-references: UNIPROT:Q8XBZ6; GB:BA000007; PIDN:BAB34581.1; PID:g13360618; GSPDB:G
C:Genetics:
A:Experimental source: strain O157:H7, substrain RMD 050952
A:Gene: ECE1158

Alignment Scores:
Pred No.: 1.84e-36 Length: 413
Score: 545.00 Matches: 135
Percent Similarity: 50.12% Conservative: 76
Best Local Similarity: 32.07% Mismatch: 180
Query Match: 23.20% Indels: 30
Gaps: 2
DB:

US-10-034-985-1 (1-1323) x F90773 (1-413)

Qy 37 ATTCCGTTAAATCCCGCAATCTCGTTCGTCAGAGTGAGCCGGAG---CTGAAGCTGGAA 93
Db 13 ValValLeuLeuAlaSerAsnAlaGlnAlaGlnThrValProGluGlyTyrGlnLeuGln 32
Qy 94 AGTGTGCTGATTCGTCTAGTCTGTCATGTCGTGCTGCCA---ACCAAGGCCACCAACTG 150
Db 33 GlnValLeuMetSerArgHisAsnLeuArgAlaProLeuAlaAsnAsnGlySerVal 52
Qy 151 ATCAGGATGTCACCCAGACGATGCGCAACCTGCGCGGTAATAAATCTGGGTTGGCTGACA 210
Db 53 LeuGlnLysSerThrProAsnLysTrpProGluTrpAspValProGlyGlyGlnLeuThr 72
Qy 211 CCGCGNGTGTGAGCTAATCTCGCTATCTCGGCATTTACCAACGCCAGCGCTGCTGAGCC 270
Db 73 ThrLysGlyGlyValLeuGluValTyrMetGlyHisTyrMetArgGluTrpLeuAlaGlu 92
Qy 271 GACGATTCGTGCGGAAAGGCTGCCCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGAT 330
Db 93 GlnGlyMetValLysSerGlyGluCysProProAspThrValTyrAlaTyrAlaAsn 112
Qy 331 GTCAGCAGGCTACCCGTAAACAGCGAGACCTTCGCCCGCGGCTGGCACCTGACTGT 390
Db 113 SerLeuGlnArgThrValAlaThrAlaGlnPhePheIleThrGlyAlaPheProGlyCys 132
Qy 391 GCATAACCGTACATACCCAGGCGAGATACGTCAGTCCCGCTCCGCTTATTTAATCTCTA 450
Db 133 AspIleProValHisHisGlnGluLysMetGlyThrMetAspProThrPheAsnProVal 152
Qy 451 AAAACTGGCTTTGCCAATCGGATAACGGAACGTCGACTGACGCGATCTCCAGCAGGCA 510
Db 153 IleThrAsp-----AspSerAlaAlaPheSerGluGlnAlaValAlaAlaMet 168
Qy 511 GGAGGTCAATTCGCTGACTTACCGGCGCATCGCAACGCGCTTCGCGAATCGGAACGG 570
Db 169 GluLysGluLeuSerLysLeu-----GlnLeuThrAspSerTyrGlnLeuLeuGluLys 186
Qy 571 GTGCTTAATTTTCGCAACTCAAACTTGTGCTTAAACGTCGAGAAACAGACGAAAGCTGT 630
Db 187 IleValAsnTyrLysAspSerProAlaCys-----LysGluLysGlnGlnCys 202
Qy 631 TCATAACGCGCATTTACCATCGCACTCAAGTGAAGCCGCGCAGCAATGTCTCATTAACC 690
Db 203 SerLeuValAspGlyLysAsnThrPheSerAlaLysTyrGlnGlnGluProGlyValSer 222
Qy 691 GGTGCGGTAAAGCTCGCATCAATGCTGACGAGATATTTCTCTGCAACAGCACAGGA 750
Db 223 GlyProLeuLysValGlyAsnSerLeuValAspAlaPheThrLeuGlnTyrTyrGluGly 242
Qy 751 ATGCGC-----GAGCGGGTGGGAAGGATCACCAGTTTACACACCATGGAACACCTTG 804
Db 243 PheProMetAspGlnValAlaTrpGlyGluLysSerAspGlnGlnTrpLysValLeu 262
Qy 805 CTAAGTTTGTATACCGCGCAATTTTATTTGCTCAACAGCAGCGCAGAGTTGCCCCGAGC 864
Db 263 SerLysLeuLysAsnGlyTyrGlnAspSerLeuPheThrSerProGluValAlaArgAsn 282
Qy 865 CGCGCCACCCCGTTTATTTGATTTGATCATGCGAGCTTGAGCCGCCATCCACCGCAAAA 924
Db 283 ValAlaLysProLeuValSerTyrIleAspLysAlaLeu-----295
Qy 925 CAGCGGTATGCTGACATTACCCACTTCA-----GTACTGTTTATTTCGCGGACAC 975
Db 296 -----ValThrAspArgThrSerAlaProLysIleThrValLeuValGlyHis 311
Qy 976 GATACTAATCTGCAATCTCGCGCGCACTGGAGCTCAAC---TGGACGCTTCCCGGT 1032
Db 312 AspSerAsnIleAlaSerLeuLeuThrAlaLeuAspPheLysProTyrGlnLeuHisAsp 331
Qy 1033 CAGCCGATACACGCGCGCAGGTGTGTAAGTGTGTTTGAACGCTCGGCTCGCTAAGC 1092

```

Db      332 GlnAsnGluArgThrProIleGlyGlyLeuValPheGlnArgTrpArgSerLys 351
QY      1093 GATACAGCCAGTCGATTCAGTTTCCTGCTTCAGACTTTACAGCAGATCGGTGAT 1152
Db      352 AlaAsnArgAspLeuMetLysIleGluTyrValTyrGlnSerAlaGluGlnLeuArgAsn 371
QY      1153 AAAACCGCGTGTCTATTAAATACCGCCCGGAGAGGTGAACCTGACCTCGGAGGATGT 1212
Db      372 AlaAspAlaLeuThrLeuGlnAlaProAlaGlnArgValThrLeuGluLeuSerGlyCys 391
QY      1213 GAAGAGCGAAATGCGCAGGCGCATGTTCGTTCGTCAGCGTTTTCAGCAAAATCGTGAATGAA 1272
Db      392 ---ProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerValLeuAsnGlu 410
QY      1273 GCA 1275
Db      411 Ala 411

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: B85636
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: B85636
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-413 <STO>
A/Cross-references: UNIPROT:Q8XBZ6; GB:AE005174; NID:gl2514273; PIDN:AAG55550.1; GSPDB:Q
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: agp

Alignment Scores:
Pred. No.: 1.84e-36 Length: 413
Score: 545.00 Matches: 135
Percent Similarity: 50.12% Conservativity: 76
Best Local Similarity: 32.07% Mismatches: 180
Query Match: 23.20% Indels: 30
DB: 2 Gaps: 10

US-10-034-985-1 (1-1323) x B85636 (1-413)
QY      37 ATTCGGTTAAACCCGCAATCTGCATTCGCTCAGAGTCAGCGGAG---CTGAAGCTGGAA 93
Db      13 ValValLeuLeuAlaSerAsnAlaGlnAlaGlnThrValProGluGlyTyrGlnLeuGln 32
QY      94 AGTGTGTGATGTCAGTCGTGTCATGTCGTGTCCTCCA---ACCAAGGCCAGCAACTG 150
Db      33 GlnValLeuMetSerArgHisAsnLeuArgAlaProLeuAlaAsnAsnGlySerVal 52
QY      151 ATGCAGATGTCACCCAGACGCATGACCAACCTGCGCGGTAAACCTGGGTGCTGACA 210
Db      53 LeuGluGlnSerThrProAsnLysThrProGluThrAspValProGlyGlyGlnLeuThr 72
QY      211 CCGCGNGTGTGAGTAAATCGCTATCTCGGACATTACCAACGCCAGCGTCTGGTAGCC 270
Db      73 ThrLysGlyValLeuGluValTyrMetGlyHisTyrMetArgGluThrLeuAlaGlu 92
QY      271 GACGATTTGTCGCGCAAAAGGCTGCCCGCAGTCTGTCAGTTCGCGATTTATTCGTGAT 330
Db      93 GlnGlyMetValLysSerGlyGluCysProProAspThrValTyrAlaTyrAlaAsn 112
QY      331 GTCGACGAGTACCCGTAACAGCGGAGACCTTCGCCCGCGCTGGCACCTGACTGT 390
Db      113 SerLeuGlnThrValAlaThrAlaGlnPhePheIleThrGlyAlaPheProGlyCys 132
QY      391 GCATATACCGGTACATACCCAGGCAGATACGTCCAGTCCCGATCCGTTTATTTAATCCTCTA 450

```

RESULT 7

S25627

glucose-1-phosphatase (BC 3.1.3.10) precursor - Providencia rettgeri

C/Species: Providencia rettgeri

C/Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

A/Accession: S25627

R/Ricci, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.

submitted to the EMBL Data Library, September 1992

A/Reference number: S25627

A/Accession: S25627

```

Db      133 AspIleProValHisIleGlnGluLysMetGlyThrMetAspProThrPheAsnProVal 152
QY      451 AAAACTGGCGCTTTGCCAACTGAGATACGCGAAGCTGACTACGCGATCCTCAGCAGGGCA 510
Db      153 IleThrAsp-----AspSerAlaAlaPheSerGluGlnAlaValAlaMet 168
QY      511 GGAGGTCAATTCGTGACTTTACCGGGCATCGGCAACGGCGTTTCGCAACTGGAAACGG 570
Db      169 GluLysGluLeuSerLysLeu-----GlnLeuThrAspSerTyrGlnLeuLeuGluLys 186
QY      571 GTCTTAATTTTCGCAACTTGTGCTTAAACGTCGAGAAACAGCAGCAGAAAGCTGT 630
Db      187 IleValAsnTyrLysAspSerProAlaCys-----LysGluLysGlnGlnCys 202
QY      631 TCATTAAAGCAGGCATTACATCGGAACCTCAAGGTGAGCGCCGACAAATGTCTCATTAAC 690
Db      203 SerLeuValAspGlyLysAsnThrPheSerAlaLysTyrGlnGlnGluProGlyValSer 222
QY      691 GGTGCGGTAAAGCTCGCATCAATGCTGCGAGATATTTCTCTGCAACAAAGCAGCAGGGA 750
Db      223 GlyProLeuLysValGlyAsnSerLeuValAspAlaPheThrLeuGlnTyrTyrGluGly 242
QY      751 ATGCCG-----GAGCGCGGTGGGAAAGGATCACCAGATTTCACACAGTGGAAACACCTTG 804
Db      243 PheProMetAspGlnValAlaTyrGlyGluLysSerAspGlnGlnTyrLysValLeu 262
QY      805 CTAAGTTTGCATTAACCGGCANNTTATTCGTCACACGCGCAGAGGTTGCCGCGAGC 864
Db      263 SerLysLeuLysAsnGlyTyrGlnAspSerLysPheThrSerProGluValAlaArgAsn 282
QY      865 CGCGCCACCCCGTTATTGGATTTCATCGGACGCTTGACGCCCATCCACCGCAAAA 924
Db      283 ValAlaLysProLeuValSerTyrIleAspLysAlaLeu----- 295
QY      925 CAGCGGTATGCTGTGACATTACCCACTTCA-----GTACTGTTTATTCGCGGACAC 975
Db      296 -----ValThrAspArgThrSerAlaProLysIleThrValLeuValGlyHis 311
QY      976 GATACTAATCTGCAATCTCGCGGCGGACCTGAGCTCAAC---TGGACGCTTCCCGGT 1032
Db      312 AspSerAsnIleAlaSerLeuThrAlaLeuAspPheLysProTyrGlnLeuHisAsp 331
QY      1033 CAGCGGATAAACCGCGCGCAGCTGTGTAACCTGCTGTTTCAACGCTGCGCTCGCTAAGC 1092
Db      332 GlnAsnGluArgThrProIleGlyLysIleValPheGlnArgTrpArgSerLys 351
QY      1093 GATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCAGACTTTACAGCAGATCGGTGAT 1152
Db      352 AlaAsnArgAspLeuMetLysIleGluTyrValTyrGlnSerAlaGluGlnLeuArgAsn 371
QY      1153 AAAACCGCGTGTCTAATAATACGCGCGCGGAGAGGTGAACCTGACCTCGCAGGATGT 1212
Db      372 AlaAspAlaLeuThrLeuGlnAlaProAlaGlnArgValThrLeuGluLeuSerGlyCys 391
QY      1213 GAAGCGCAATCGCAGGCGCATGTGTCGTCGAGGTTTACGCCAAATCGTGAATGAA 1272
Db      392 ---ProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerValLeuAsnGlu 410
QY      1273 GCA 1275
Db      411 Ala 411

```

A:Molecule type: DNA
A:Residues: 1-417 <RIC>
A:Cross-references: UNIPROT:Q52309; EMBL:X68201; NID:g45771; PIDN:CAA48288.1; PID:g45772
A:Experimental source: strain PV7
C:Genetics:
A:Gene: agp
C:Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hydrolysis
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-417/Product: glucose-1-phosphatase #status predicted <MAT>
F:42/Active site: His (phosphohistidine intermediate) #status predicted
F:314/Active site: His #status predicted

Alignment Scores:
Pred. No.: 3,92e-36 Length: 417
Score: 541.00 Matches: 129
Percent Similarity: 50.35% Conservative: 84
Best Local Similarity: 30.50% Mismatches: 182
Query Match: 23.03% Indels: 28
DB: 2 Gaps: 9

US-10-034-985-1 (1-1323) x S25627 (1-417)

Qy	28	TCCTCTCTGATTCCGTTAACCCGCAATCGCATTCGCTCAGAGTGAGCGGAGCTGAAG	87
Db	13	AlaLeuPheAlaProIleAlaProThrMetAlaSerThrAspAsnGlnAlaAspMetVal	32
Qy	88	CTGGAAAGTGTGGTGATTGTTCAGTCGTCATGTGTGTGCTCTCAACCAAGGCACGC	147
Db	33	LeuAspGlnValLeuValLeuSerArgHisAsnLeuArgThrProIleValAsnThrGly	52
Qy	148	CTGATCGAGATGTCACCCAGACGCATGCCCAACTCGCGGTAAACCTGGGTTGGCTG	207
Db	53	IleLeuThrGluValThrAspLysLysTrpAspTrpAspAlaLysSerGlyTyrLeu	72
Qy	208	ACACCGCGGTGGTGAGCTAATCGCTATCTCGGACATTAACCAAGCCAGCGCTCGTGA	267
Db	73	ThrThrGlnGlyAlaLeuGluValTyrMetGlyHisTyrPheArgGluTrpIleAsp	92
Qy	268	GCCGACGGATTGTGGCGAAAGGGCTGCCGCGAGTCGTGTCAG--GTCGCGATTATT	324
Db	93	GlnAsnLysLeuLeuAlaAspGluLeuCysProThrSerAsnGluAspIleTyrLeuTyr	112
Qy	325	GCTGATGTCGACGAGCTACCCGTAAACAGGCGAAGCCTTCGCCCGCGGTGGCACCT	384
Db	113	ThrAsnSerLeuGlnArgThrIleAlaThrAlaGlnPheAlaAlaGlyAlaPhePro	132
Qy	385	GACTGTGCAATACCGTACATACCCAGGAGATACGTCCAGTCCCGATCCGTTATTAAT	444
Db	133	GlyCysLysValAsnIleHisHisGlnProGluIleGlyLysMetAspProValPheAsn	152
Qy	445	CCTCTAAAACTGGCGTTTCCCACTGGATAACCGCAACGTGACTGACGCGATCCTCAGC	504
Db	153	ProIleLeuThr-----AsnGlySerProGluPheLysGln	164
Qy	505	AGGCGAGGAGGTCATTTGCTGACTTTTACCGGG-----CATCGCAACCGGTTTCGC	558
Db	165	LysAlaLeuAlaAlaMetAspTyrTyrLeuLysGlyLeuSerLeuLysAlaGlyTyrGlu	184
Qy	559	GAACCTGAACGGGTGCTTAATTTTCCGCAATCAAACTTTGTGCCTTAAACGTGAGAACAG	618
Db	185	GluLeuAspThrValLeuAsnIleLysAspSerGlnLysCys-----LysThr	200
Qy	619	GAGCAAGCTGTTTCATTAACGACGAGCATTAACCATCGAACTCAAGGTGAGCGCGACAAT	678
Db	201	AspLysLeuCysAsnLeuAspSerGlnLysAsnSerPheIleLeuAlaAspLysGlu	220
Qy	679	GTCTCATTAACCGGTGGTAAAGCTCCGATCAATGCTGACGGAGATATTTCTCTGCCAA	738
Db	221	ProGlyValSerGlyProLeuLysIleAlaAsnSerAlaValAspAlaIleAspLeuGln	240
Qy	739	CAAGCACAGGAATGCCG-----GAGCGGGGTGGGGAAGGATCAACCGATTCAACACGAG	792
Db	241	TyrTyrGluGlyPheProAlaAspGlnValAlaTrpGlyLeuValAspThrProGluLys	260

[illegible][illegible]


```

Db 81 ArgHisTyrAlaAlaGlnGlyLeuLeuYleProGlyAspCys-----AlaSerVal 97
Qy 316 GCGATTATTGCTGATGTCGACGAGCTACCGTAAACAGCGGAGCGGCTTCGCGCGCGG 375
Db 98 TyrAlaTrpAlaAsnValThrGlnArgThrIleAlaThrAlaYleAlaTyrArgGluThr 117
Qy 376 CTGCGACCTGACTGTGCAATAACCGTACATACCAGGAGATAGTCCAGTCCGATCCG 435
Db 118 LeuAlaProGlyCysProValThrValAsnThrValGlyGlu---GlyAsnIleAspPro 136
Qy 436 TTATTTAATCTCTAATAACCTGGCTTGGCACTGATTAACGCGAAGCTGACTGACGCG 495
Db 137 MetPheGluProValYleAlaGlyIleValYleAlaAspHisAlaLeuAlaAa-gAlaAla 156
Qy 496 ATCTCTACGAGGCGAGCGGTGCTCAATTGCTGACTTTTACCGGCGATCGCAACGCGCTT 555
Db 157 ValAlaGlyArgValGlyGlyAspLeuThrAlaTrpSerAlaSerHisAsnGlnGluAla 176
Qy 556 CGGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTCCTT----- 603
Db 177 GluGlnLeuAspAlaLeuLeuMetGlnCysAspYleGlyProCysProProAlaProGly 196
Qy 604 AAGCTGAGAAACGAGGAGAAAGCTGTTCATTACGCGAGCATTACCATCGGAACCTAAG 663
Db 197 LysArgArgValPheAspAlaLysProGlyPheValAspGly-----GluGluLeu--- 213
Qy 664 GTGAGCGCCGACATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACGGAG 723
Db 214 -----AlaGlyLeuSerGlyProGluAlaPheAlaSerGlyValThrGlu 228
Qy 724 ATATTTCTCTGCAACAGACAGGGAATCGCGAGCGGGGTGGGA---AGGATCAC 780
Db 229 SerLeuLeuMetAlaTrpAlaAspGlyArgAspPheAlaGlyLeuGlyTyrLysSerLeu 248
Qy 781 GATTACACAGGTGGAACACTGCTAAGTTTGATTAACCGGCAATTTTATTGCTACAA 840
Db 249 AspGluGluAlaLeuThrArgSerPhePheLeuHisGlnAlaGlyPheAspLeuArgLeu 268
Qy 841 CGCAGCGCAGAGGTGCGCGAGCGGCGGCGGCTTATTTGATTTGATTCATGGCAGCG 900
Db 269 ArgThrProTyrValAlaAlaGlnThrLeuAlaGlyHisLeuAlaAspArgLeuAlaAlaThr 288
Qy 901 TTGACGCGCCCATCCACGCGAAAAACAGCGGTATGCTGATTCATTCACCTTCAGTACTG 960
Db 289 LeuArg-----AspGlyAlaAlaAlaIleGly---ProValAspAlaArgLeuVal 304
Qy 961 TTTATTCGCGACGATACTAATCTGCGAAATCTCGCGCGGCGACCTGGAGCTCAACTGG 1020
Db 305 IleIleAlaGlyHisAspGlyThrLeuAlaSerLeuGlyGlyLeuLeuArgMetGluTrp 324
Qy 1021 ACGTCTCCCGGT---CAGCGCGGATAACACGCGCGGTGCTGAACCTGGGTGTTGAACGC 1077
Db 325 ThrLeuProGlyTyrGlnProAsnGlnIleGlnProGlyAlaLeuValPheGluArg 344
Qy 1078 TGGCGTCCGCTAAGCGATAACAGCAGTGGATTGAGTTTCGCTGCTTCCTCCAGCTTTA 1137
Db 345 TrpArgArg---AspAspGlyValArgValValArgValArgPheThrGlyGlnSerLeu 363
Qy 1138 CAGCAGATGGGTGATAAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTG 1197
Db 364 SerGlnLeuArgAsnMetThrAlaLeuAspAlaLysThrProProLeuSerAlaProVal 383
Qy 1198 ACCCTGCGAGATCTGAAGAGCGAAATCGCAGCGCATGTGCTGTTGGCAGGTTTACG 1257
Db 384 PheValGlnGlyCysGlyThrAlaThrProAlaPheAspCysArgLeuGluAspPheGlu 403
Qy 1258 CAAATCGTGAATGAAGCA 1275
Db 404 ThrValValArgGlyAla 409

```

RESULT 10
AG0632

glucose-1-phosphatase precursor (G1Pase), secreted [imported] - Salmonella enterica subsp.
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0632
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08242.1; PID:gl6502289; GSPDB:GN00176
C;Genetics:
A;Gene: STY1153

Alignment Scores:

Pred. No.: 9,646-35 Length: 413
Score: 524.00 Matches: 135
Percent Similarity: 48.90% Conservative: 65
Best Local Similarity: 33.01% Mismatches: 185
Query Match: 22.31% Indels: 24
DB: 2 Gaps: 10

US-10-034-985-1 (1-1323) x AG0632 (1-413)

```

Qy 64 GTCAGAGTGTAGCCGAG---CTGAAGCTGGAAGTGTGTGTATGTCTAGTCATCGT 120
Db 22 AlaGlnThrThrProGluGlyTyrGlnLeuGlnValLeuMetSerArgHisAsn 41
Qy 121 GTCGCTGCTCCA---ACCAAGGCCGCACTCATCGAGATGTCACCCAGCAGCTG 177
Db 42 LeuArgAlaProLeuAlaAsnAsnGlyAsnValAlaGlnSerThrProAsnAlaTrp 61
Qy 178 CCAACCTGCGCGGTAACAACTGGGTGTGTCACCGCGGTGTGAGCTTAATCGCTAT 237
Db 62 ProAlaTrpAspValProGlyGlyGlnLeuThrThryGlyGlyValLeuGluValTyr 81
Qy 238 CTGCGACATTACCAACGCCAGCGCTCTGTAGCCGAGCGGATGCTGGCGAAAAAGGGTGC 297
Db 82 MetGlyHisTyrThrArgGluTrpLeuValAlaGlnGlyLeuIleProSerGlyGluCys 101
Qy 298 CGCAGCTCTGTCAGTCCGATTTATGCTGATGTCGACGAGCGTACCCGTAAACAGGC 357
Db 102 ProAlaProAspThrValTyrAlaTyrAlaAsnSerLeuGlnArgThrValAlaThrAla 121
Qy 358 GAAGCTCTCGCCCGCGGTGCGACCTGACTGTGCAATACCGTATACCCAGGCAGAT 417
Db 122 GlnPheIleThrSerAlaPheProGlyCysAspIleProValHisHisGlnGluLys 141
Qy 418 ACCTCCAGTCCCGATCCGTTTATTAATCTCTTAAACCTGGCGTTTGGCCAACTGGATAAC 477
Db 142 MetGlyThrMetAspProThrPheAsnProValIleThrAspAspSerAlaAlaPheArg 161
Qy 478 GCGAAGCTGACTCAGCGGATCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGG 537
Db 162 GlnGlnAlaValGlnAla---MetGluLysAlaArgSerGlnLeu----- 175
Qy 538 CATCGGCAAAACGGGTTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTG 597
Db 176 HisLeuAspGluSerTyrLysLeuLeuGlnIleThrHisTyrGlnAspSerProSer 195
Qy 598 TGCCTTAAACGTGAGAAACAGGAGAAAGCTGTTTATTAAACGCGACGATTCACCATCGGAA 657
Db 196 Cys-----LysGluLysHisGlnCysSerLeuIleAspAlaLysAspThrPhe 211
Qy 658 CTCAGGTGAGCGCCGACCAATGCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTG 717
Db 212 SerAlaAsnTyrGlnGlnGluProGlyValGlnGlyProLeuLysValGlyAsnSerLeu 231

```



```

Db      222 ProTrpAlaSerProGlnThrMetGlnArgLeuSerArgLeuTyrAspPheSerPheArg 241
Qy      823 CAATTTTATTTGCTACAAACGACGCGAGGTTCCCGCGAGCGCGCCACCCGCTTATTG 882
Db      242 PheLeuPheGlyIleTyrGlnGlnAlaGlyLysAlaArgLeuGlnGlyValLeuLeu 261
Qy      883 GATTGTGATCATGGCAGCGTTGACGCCCATCCACCGCAAAACAGGCGGTATGTGTGACA 942
Db      262 AlaGlnIleArgLysAsnLeuThrLeuMetAlaThrThrSerGln----- 276
Qy      943 TTACCCACTTCAGTACTGTTTATTGCGGACACAGTACTAATCTGGCAAAATCTCGGCGGC 1002
Db      277 LeuPro---LysLeuLeuValTyrSerAlaHisAspThrThrLeuValAlaLeuGlnMet 295
Qy      1003 GCACTGGAGTCAACTGGACGCTTCCCGGTGAGCGGATACACGCGCGAGGTGTGAA 1062
Db      296 AlaLeuAspValTyr-----AsnGlyGlnGlnAlaProTyrAlaSerCys 310
Qy      1063 CTGGTGTGTTGAACGCTGGCGTAAAGCGATTAACAGCCAGTGGATTTCAGGTTTCGCTG 1122
Db      311 HisIlePheGluLeuTyrGlnGluAspSerGlyAsn-----PheSerValGluMet 327
Qy      1123 GTCTTCAGACTTTACAGCAGATCGGTGATAAAACGCGCTGTCTAATTAATACGCCGCC 1182
Db      328 TyrPhe-----ArgAsnGluSerAspLysAlaProTyrProLeuSerLeuPro--- 343
Qy      1183 GGAGAGGTGAACGCTGCGGAGGATGTGAAGAGCGAAATGGCGGCGCATGTGTTCG 1242
Db      344 -----GlyCysProHisArg-----CysPro 350
Qy      1243 TTGCAGGTTTACGCAAAATCGTAATGAAGACGCGATACCG-----GCG 1287
Db      351 LeuGlnAspPheLeuArgLeu---ThrGluProValValProLysAspTyrGlnGlnGlu 369
Qy      1288 TGCAGTTTGGATCT 1302
Db      370 CysGlnLeuAlaSer 374

RESULT 12
T16058
hypothetical protein F13D11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16058
R:Rulton, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F13D11.
A:Reference number: S69020
A:Accession: T16058
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-416 <FUL>
A:Cross-references: UNIPROT:Q19390; EMBL:U40939; NID:g1072175; PID:g1072178; PIDN:AAA817
C:Genetics:
A:Introns: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1
C:Superfamily: mammalian acid phosphatase

Alignment Scores:
Pred. No.: 0.00968 Length: 416
Score: 133.00 Matches: 104
Percent Similarity: 36.94% Conservative: 60
Best Local Similarity: 23.42% Mismatches: 174
Query Match: 5.66% Indels: 106
DB: 2 Gaps: 24

US-10-034-985-1 (1-1323) x T16058 (1-416)

Qy      1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTTCGGTTAACCCGCAATCTGCA 60
Db      1 MetArgLeuValLeuLeuPhe-----PheLeuPheProVal-----Ala 14

```

```

Qy      61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGTGATGTGTGATGTGTGATGTGTGAT 120
Db      15 PheGly-----LysLeuPheValGlnIleTyrPheArgHisGly 28
Qy      121. GTCGGTCTCCAAACCAAGGCCACGCAACTGATCAGGATGTCAACCCAGAC----- 171
Db      29 GluArgThrProGlyHisTyrLeuTyrPheProGlyAspAspLeuAsnValAspTyr 48
Qy      172 -----GCATGGCCCAACTGCGCGGTAAACTGGTTGGCTGACACCGCGNGGTGGT 222
Db      49 GlnGlnIleAlaTyrPro-----GlyGluLeuThrLysArgGlyLe 62
Qy      223 GAGCTAATCCCTATCTCGGACATTTACAAAGCCAGCGTCTGTGTAGCCGAGGATTGCTG 282
Db      63 LeuGluGluPheGlnLeuGly-----GlnArgLeuArgLysIleTyrGlyGluHisPhe 80
Qy      283 GCGAAAAAGGCTGCCCGAGTCTGTGTGAGTCCGATTTATGCTGATGTGTGAGCAGCGT 342
Db      81 GlyAspThrTyrGlnProArgAspPheHisVal-----TyrThrGlyLysAspAsnArg 98
Qy      343 ACCCGTAAACAGCGCAAGCCTTCGCGCGGTGGCGCTGGCACCTGACTGTGCAATAACCGTA 402
Db      99 ThrSerAlaSerAlaGlnAlaMetPheAlaGlyPheLeuProPro----- 113
Qy      403 CATACCCAGCAGATAGTCCGATCCGATCCGATCCGATTTAATCTCTAAAACATGGCGTT 462
Db      114 ---AsnGluAspGlnThrTrpAsnTyrGluLeuLysTrpGlnPro-----Val 128
Qy      463 TGCCAACTGGATAACGCAACGTGACGCGATCCTCAGCAGGCGGAGGAGGTCAT 522
Db      129 AlaGlnLeuThrAspGluSerIleAspTyrValSerLeu-----GlyAlaIle 144
Qy      523 -----GCTGACTTTACCGGCACTCGCAACCGCGTTTCGCGAACTGGACCGGTGCTT 576
Db      145 AspAsnCysProValTyrGlyGluAlaGlnArgLysSerSerGluTyrAlaGluValMet 164
Qy      577 AATTTCCGCAATCAAACTTGTGCTTAAACGTGAGAAACAGACGAAAGCTGTTCATTA 636
Db      165 Asp-----GlnMetGluLysTyrAsp-----AlaGluLeu 174
Qy      637 ACCGAGCATTTACATCGGNACTCAAGTGAGCGCGCACTGTCTCATTAACCGGTGCG 696
Db      175 LeuGlnLeuValArgAsnHis-----AlaAspGlu---ProIleValGluAla 189
Qy      697 GTAAGCCTCGCATCAATGCTGCGGAGATA-----TTTCTCTGCAACAGCACAG 747
Db      190 ValLysTyrAsnHisValIleAspSerLeuLysValArgTyrIleLeuGlnAspAspArg 209
Qy      748 GGAATCCCGAGCGCGGTGGGAAAGATCACCGATTACACAGGTGGAACACCTTGCTA 807
Db      210 ---LeuProTyrProGluTyrAlaArgGlyTyrGluAsnArgIleLeuAsnMetSerPhe 228
Qy      808 AGTTTGCATACCGGCAATTTTATTGTCTACACGCGCAGGTTGCCCGCAGCGCG 867
Db      229 LeuIleHisAspAla-----ValValLysValGln 238
Qy      868 GCCACCCCGTTATTGGAT-----TTGATCATGCGCGGTGTGACGCCCATCCA 915
Db      239 AsnAspSerValGlyAspTyrHisAsnGluLeuValMetSerTyrPheGluThrHisLeu 258
Qy      916 CCGCAAAACAGCGGTATGTTGTGACATTATCCCACTTCAGTACTGTTTATTTGCGGACAC 975
Db      259 GlnLysAsnSerThrLysGlyVal-----PheIleSerGlyHis 271
Qy      976 GATACTAATCTGCAAAATCTCGCGGCGGCACTGAGCTCAACTGGAGCTTCCCGGTGAG 1035
Db      272 AspThrAsnLeuValThrIleTyrGluSerLeuArgLeuAsp-----GlyHis 287
Qy      1036 CCGGATAACACCGCGCA-----GGTGTGAACTGTGTTCGAAACGCTGGCGTGGCTAAGC 1092
Db      288 ProGluAspIleProAsnTyrGlyAlaHisIleAlaIleGlu----- 301
Qy      1093 GATAACAGCCAGTGGATTTCAGGTTTCGCTGCTCTTCAGACTTTTACAGCAGATGCGTGAT 1152

```


Db 210 CysGluClnThrHisSgLyLeuLeuLeuProTrpAlaSerProGlnThrValGlnArg 229

Qy 793 TGGAAACACTTGTCTAAGTTTGTCATAACGCCCAATTTATTATTGCTACTAACCGCACGACAG 852
::: ||| :::
Db 230 LeuSerGlnLeuLysAspPheSerPheLeuPheLeuGlyIleHisGluGlnValGln 249

Qy 853 GTTCCCGCAGCGCCGCCACCCTGTTATTCGATTGTGATCATGGCAGCGTTGCAGCCCCCAT 912
||||| ::: ||||| |||||
Db 250 LysAlaArgLeuGlnGlyValLeuLeuAlaGlnIleLeuLysAsnLeuThrLeuMet 269

Qy 913 CCACCGCAAAAAACGGCGTATGTGTGCATTAACCCACTTCAGTACTCTTTATTGCCCAGA 972
||| ||| ::|||
Db 270 AlaThrThrSerGlnPhe-----ProllysLeuLeuValTySerAla 283

Qy 973 CACGACTACTAATCTGGCAAATCTCGCGCGCCTACTGGAGCTCACTGGACGCTTCCCGGT 1032
||||| ||| ||| ||||| ::
Db 284 HisAspThrThrLeuValAlaLeuGlnMetAlaAsnValTy----- 298

Qy 1033 CAGCCGGATTAACAGCGCCAGGTGGTGAACCTGGTGTTCGAACGCTGGCGTCGGCTAAGC 1092
||| ||| ::||| |||||
Db 299 AsnGlyLysGlnAlaProTyAlaSerCysHisIlePheGluLeuTyrGln-----Glu 316

Qy 1093 GATAACAGCCAGTGGATTTCAGGTTTCGCTGCTCTCCAGACTTTACAGCAGATGCGTGAT 1152
||||| ||| ||| ::||| |||||
Db 317 AspAsnGlyAsnPhe---SerValGluMetTyPheArgAsnAspSerLys----- 332

Qy 1153 AAAACGCGCTGTCTATTAAATACGCCCGCGAGAGGTGAACTGACCCCTGGCAGGATGT 1212
||| ||| ||| ||| |||||
Db 333 LysAlaProTrpProLeuIleLeuPro-----GlyCys 343

Qy 1213 GAAGAGCGAAATCGCAGGCGCATGTGTTCTGTCAGGTTTTTACGCAAAATCGTGAATGAA 1272
||| ||| ||| ||| |||||
Db 344 ProHisArg-----CysProLeuGlnAspPheLeuArgLeu---ThrGlu 357

Qy 1273 GCACGCATACCG 1284
||| |||||
Db 358 ProValIlePro 361

RESULT 15

JH0610

A;Title: phosphatase (EC 3.1.3.2) ACP precursor [validated] - human

N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphatase

C;Species: Homo sapiens (man)

C;Date: 17-Aug-1992 #sequence revision 01-Dec-1995 #text change 09-Jul-2004

S;Accession: JH0610; JS0693; A38608; S01331; A32419; S11147; S38863; S41251; S11147

R;Sharief, F.S.; Li, S.S.L.

Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992

A;Title: Structure of human prostatic acid phosphatase gene.

A;Reference number: JH0610; MUID:92272747; PMID:1375464

A;Accession: JH0610

A;Molecule type: DNA

A;Residues: 1-386 <SHA>

A;Cross-references: UNIPROT:P15309; GB:M97580; GB:M97581; GB:M97582; GB:M97583; A38608

A;Molecule type: mRNA

A;Residues: 1-386 <SH3>

A;Cross-references: GB:M97589; NID:g189611; PIDN:AAA60021.1; PID:g189613

R;Van Btten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.

J. Biol. Chem. 266, 2313-2319 1991

A;Title: Covalent structure, disulfide bonding, and identification of reactive site

A;Reference number: A38608; MUID:91115848; PMID:1989985

A;Accession: A38608

A;Molecule type: mRNA

A;Residues: 1-386 <VAN>

A;Cross-references: GB:M34840; NID:g189620; PIDN:AAA69694.1; PID:g189621

A;Note: part of this sequence, including the amino end of the mature protein, w

R;Vihko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Huhtala, M.L.

FEBs Lett. 236, 275-281, 1988

A;Title: Molecular cloning and sequence analysis of cDNA encoding human prost

A;Reference number: S01331; MUID:88312981; PMID:2842184

A;Accession: S01331

A;Molecule type: mRNA

A;Residues: 1-14,'A','16,'ASC',20,'CF',23,'C',25-65,'WIWPTHPA',74-211,'A',213-388

A; Cross-references: EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID:g28322
A; Note: part of this sequence, including the amino end of the mature protein, was confirmed by R; Sharief, F.S.; Lee, H.; Leuderman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S.S. Biochem. Biophys. Res. Commun. 160, 79-86, 1989
A; Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequencing
A; Reference number: A32419; MUID:89228054; PMID:2712834
A; Accession: A32419
A; Molecule type: mRNA
A; Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-94, 'D', 96-115, 'R', 117-214, 'S', 216-293, 'T'
A; Cross-references: GB:M24902; NID:g189618; PIDN:AAA60022.1; PID:g189619
A; Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue 116 as Arg.
R; Tailor, P.G.; Govindan, M.V.; Patel, P.C. Nucleic Acids Res. 18, 4928, 1990
A; Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-length cDNA
A; Reference number: S11147; MUID:90370491; PMID:2395659
A; Accession: S11147
A; Status: preliminary; translation not shown
A; Molecule type: mRNA
A; Residues: 1-14, 'AFASC', 20, 'CF', 23, 'C', 25-45, 'H', 47-65, 'RIWTHPA', 74-138, 'E', 140-156, 'R'
A; Cross-references: EMBL:X53605; NID:g35683; PIDN:CAA37673.1; PID:g35684
R; Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W. submitted to the EMBL Data Library, April 1993
A; Description: Characterization of the promoter of the human prostatic acid phosphatase
A; Reference number: S38863
A; Accession: S38863
A; Molecule type: DNA
A; Residues: 1-40 <BAN>
A; Cross-references: EMBL:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R; Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taaviteinen, P.; submitted to the EMBL Data Library, September 1993
A; Description: Structural organization of human and rat prostate-specific acid phosphatase in the human gene promoter.
A; Reference number: S41251
A; Accession: S41251
A; Molecule type: DNA
A; Residues: 1-40 <VR>
A; Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658
R; Lee, H.; Chu, T.M.; Li, S.S.L.; Lee, C. Biochem. J. 277, 759-765, 1991
A; Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
A; Reference number: S17042; MUID:91336999; PMID:1908222
A; Accession: S17042
A; Status: preliminary
A; Molecule type: protein
A; Residues: 33-49 <LEE>
R; Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W. Biochim. Biophys. Acta 1217, 188-194, 1994
A; Title: Analysis of the promoter of the human prostatic acid phosphatase gene.
A; Reference number: S42730; MUID:94153995; PMID:8110833
A; Accession: S42730
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-40 <BA2>
A; Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R; Sharief, F.S.; Li, S.S. Biochem. Mol. Biol. Int. 33, 561-565, 1994
A; Title: Nucleotide sequence of human prostatic acid phosphatase ACPG gene.
A; Reference number: I37175; MUID:95038536; PMID:7951074
A; Accession: I37175
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 1-386 <RES>
A; Cross-references: EMBL:U07097; NID:g515995; PIDN:AAB60640.1; PID:g515997
C; Comment: This protein is synthesized under androgen regulation by epithelial cells of C;Genetics:
A; Gene: GDB:ACPP
A; Cross-references: GDB:119644; OMIM:171790
A; Map position: 3q21.3-3q25.2
A; Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
C; Function:
A; Description: catalyzes the hydrolysis of a wide range of phosphate esters
C; Superfamily: mammalian acid phosphatase
C; Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolysis

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:09:30 ; Search time 254 Seconds
(without alignments)
5334.494 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcaccatcaccatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q/cg2 1/USP0 spool_p/US10034985/runat 10032005 164240 11950/app query.fasta 1.1479
-DB=uniprot_03 -OPMT=fastan -SUFFIX=rup -MINMATCH=0 1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10034985 @CGN 1 1 271 @runat 10032005 164240 11950 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235	95.1	432	1 PPA_ECOLI	P07102 escherichia
2	2231	95.0	432	2 Q8GN88	Q8gn88 escherichia
3	2229	94.9	432	2 Q8RDK6	Q8rkd6 escherichia
4	2229	94.9	432	2 Q8RDK7	Q8rkd7 escherichia
5	2229	94.9	432	2 Q8RDK8	Q8rkd8 escherichia
6	2229	94.9	432	2 Q8RKE0	Q8rke0 escherichia
7	2228	94.8	432	2 Q8RDK9	Q8rkd9 escherichia
8	2225	94.7	432	2 Q8RDK5	Q8rkd5 escherichia
9	2222	94.6	442	2 Q83RW2	Q83rw2 shigella fl
10	2221	94.6	432	2 Q6RK08	Q6rk08 escherichia
11	2218	94.4	432	2 Q7UD08	Q7ud08 shigella fl
12	2213	94.2	446	2 Q8CWF5	Q8cw75 escherichia
13	2210	94.1	434	2 Q7AFW3	Q7afw3 escherichia
14	2210	94.1	444	2 Q8XC29	Q8xc29 escherichia
15	1374.5	58.5	433	2 Q676V7	Q676v7 citrobacter
16	1077	45.8	444	2 Q6U677	Q6u677 obesumbacte

17	1054.5	44.9	444	2	Q6TAQ8	Q6taq8 obesumbacte
18	943.5	40.2	441	2	Q669K3	Q669k3 yersinia ps
19	943.5	40.2	441	2	Q8ZFP6	Q8zfp6 yersinia pe
20	545	23.2	413	2	Q8XBZ6	Q8xbz6 escherichia
21	541	23.0	417	1	AGP_PRORE	Q52309 providencia
22	539	22.9	413	1	AGP_PRORE	P19926 escherichia
23	538	22.9	413	2	Q7UD02	Q7ud02 shigella fl
24	538	22.9	421	2	Q83RV6	Q83rv6 shigella fl
25	529	22.5	413	1	AGP_SALTY	Q39921 salmonella
26	526	22.4	414	2	Q9AAQ4	Q9aaq4 caulobacter
27	524	22.3	413	2	Q8Z7P1	Q8z7p1 salmonella
28	523	22.3	413	2	Q6EV19	Q6ev19 enterobacte
29	505	21.5	392	2	Q8PP76	Q8pp76 xanthomonas
30	505	21.5	435	2	Q8P330	Q8p330 xanthomonas
31	480.5	20.5	443	2	Q8PF53	Q8pf53 xanthomonas
32	467.5	19.9	435	2	Q6CZF4	Q6czf4 erwinia car
33	433.5	18.5	428	2	Q8GD20	Q8gd20 pseudomonas
34	392	16.7	318	2	Q8CW72	Q8cw72 escherichia
35	373	15.9	414	2	Q8PPX6	Q8ppx6 xanthomonas
36	359	15.3	421	2	Q84CN9	Q84cn9 klebsiella
37	353	15.0	421	2	Q7WSY1	Q7wsy1 klebsiella
38	233	9.9	419	2	Q8VQS2	Q8vqz2 klebsiella
39	195	8.3	426	2	Q9GNZ3	Q9gnz3 homo sapien
40	150	6.4	531	2	Q9GNZ3	Q9gnz3 leishmania
41	140	6.0	423	1	PPAL_HUMAN	P11117 homo sapien
42	140	6.0	542	2	Q9UIA2	Q9uii2 leishmania
43	139.5	5.9	333	2	Q9BZG3	Q9bzg3 homo sapien
44	139	5.9	462	2	Q7QC05	Q7qc05 anopheles g
45	137.5	5.9	406	2	Q66L09	Q66l09 xenopus lae

ALIGNMENTS

RESULT 1

PPA_ECOLI
ID PPA_ECOLI STANDARD; PRT; 432 AA.
AC P07102;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Periplasmic appA protein precursor [Includes: Phosphoanhydride
DE phosphorylase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE phytase (EC 3.1.3.26)].
GN Name=appA; OrderedLocusNames=b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-29.
RC STRAIN=K12;
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.,
RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP SEQUENCE OF 1-112 FROM N.A.
RX MEDLINE=8721766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA Touati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
acid phosphatase structural gene (appA) of *E. coli*: a negative control
of transcription mediated by cyclic AMP.";
RL Biochimie 69:215-221(1987).
RN [5]
RP SEQUENCE OF 1-17 FROM N.A.
RX STRAIN=K12;
RA MEDLINE=92049231; PubMed=1658595;
RA Dassa J., Feihi H., Marck C., Dion M., Kieffer-Bontemps M.,
Boquet P.L.;
RT "A new oxygen-regulated operon in *Escherichia coli* comprises the genes
for a putative third cytochrome oxidase and for pH 2.5 acid
phosphatase (appA)." ;
RL Mol. Gen. Genet. 229:341-352(1991).
RN [6]
RP CHARACTERIZATION, AND SEQUENCE OF 23-34.
RA Greiner R., Jany K.D.;
RT "Characterization of a phytase from *Escherichia coli*." ;
RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN [7]
RP CHARACTERIZATION, AND SEQUENCE OF 23-35.
RX MEDLINE=93256556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA Greiner R., Konietzny U., Jany K.D.;
RT "Purification and characterization of two phytases from *Escherichia coli*." ;
RL Arch. Biochem. Biophys. 303:107-113(1993).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=20161462; PubMed=10696472;
RA Glovan S., Wang G., Zhang J., Forsberg C.W.;
RT "Characterization and overproduction of the *Escherichia coli* appA
encoded bifunctional enzyme that exhibits both phytase and acid
phosphatase activities." ;
RL Can. J. Microbiol. 46:59-71(2000).
RN [9]
RP MUTAGENESIS.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase." ;
RL J. Biol. Chem. 267:22830-22836(1992).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA Lam D., Golovan S., Forsberg C.W., Jia Z.;
RT "Crystal structures of *Escherichia coli* phytase and its complex with
phytate." ;
RL Nat. Struct. Biol. 7:108-113(2000).
CC -1- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
phosphate.
CC -1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = ID-
myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: In addition to cAMP-mediated control, this enzyme is
induced when bacterial cultures reach stationary phase; its
synthesis is triggered by phosphate starvation or a shift from
aerobic to anaerobic conditions.
CC -1- SIMILARITY: Belongs to the histidine acid phosphatase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M58708; AAA72086.1; -;
DR EMBL; U00096; AAC74065.1; -;
DR EMBL; D90735; BAA35745.1; -;
DR EMBL; X05471; CAA29031.1; -;
DR EMBL; M58708; -; NOT_ANNOTATED_CDS.
DR EMBL; S63811; AAB20286.1; -;
DR PIR; B36733; B36733.
DR PDB; 1DKL; X-ray; A/B-23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR SWISS-2DPAGE; P07102; COLL.
DR EcoBASE; E80047; -;
DR EcoGene; EGI0049; appA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW 3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
KW Multifunctional enzyme; Periplasmic; Signal.
FT SIGNAL 1 22
FT CHAIN 23 432 Periplasmic appA protein.
FT ACT_SITE 39 39 Nucleophile.
FT ACT_SITE 326 326 Proton donor.
FT DISULFID 99 130
FT DISULFID 155 430
FT DISULFID 200 210
FT DISULFID 404 413
FT CONFLICT 51 66 MODVTPDAWPTWPKL -> NAGCHPRMANLAGKT (in
Ref. 3).
FT CONFLICT 75 76 EL -> DV (in Ref. 4).
FT CONFLICT 112 112 D -> S (in Ref. 4).
FT STRAND 28 38
FT STRAND 42 42
FT HELIX 49 53
FT TURN 54 54
FT TURN 55 57
FT STRAND 69 69
FT HELIX 71 90
FT TURN 91 92
FT TURN 102 103
FT STRAND 105 109
FT HELIX 113 126
FT TURN 128 129
FT STRAND 134 135
FT TURN 140 141
FT HELIX 145 147
FT TURN 149 153
FT HELIX 159 169
FT TURN 170 171
FT HELIX 174 179
FT TURN 180 181
FT HELIX 182 192
FT TURN 193 193
FT HELIX 194 196
FT HELIX 198 201
FT TURN 202 202
FT HELIX 212 215
FT STRAND 220 223
FT TURN 224 225
FT STRAND 226 229
FT HELIX 231 249
FT TURN 250 250
FT HELIX 254 257


```

FT TURN 258 258
FT HELIX 262 279
FT TURN 280 280
FT HELIX 283 301
FT STRAND 307 308
FT HELIX 310 312
FT STRAND 314 315
FT STRAND 319 324
FT HELIX 326 336
FT TURN 337 337
FT STRAND 342 342
FT TURN 343 344
FT STRAND 345 345
FT TURN 351 352
FT STRAND 354 362
FT TURN 363 366

Alignment Scores:
Pred. No.: 2,31e-172 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 1 Gaps: 0

US-10-034-985-1 (1-1323) x PPA_ECOLI (1-432)
QY 1 ATGAAGCGCATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCCGCAATCTGCA 60
DB 1 MetLysAlaIleuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTTCATGCTCATGTT 120
DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGTCCACACCGGCGCAGCACTGATCAGGATGTCAACCCAGCAGCATGGCCA 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC 240
DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTAACAACCGCAGCTCTGTGTAGCGCAGCGATTTGTCGCGGAAAAGGCGTCCCG 300
DB 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGTCCAGTCGCGATTTATGCTGATGTGACGAGCGGTACCGTAAACAGCGGAA 360
DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGGCACCTGCTGTGCAATAACCGTACATACCCAGCGCAGATACG 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTAATCTCTAAAACTGGCGTTTGCACAACTGGATAACGCG 480
DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTACCGGATCTCAGCAGGGCGAGGGGTCAATTCCTGACTTACCGGCGAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGCGCTTTCGCAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTGTGC 600
DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGGCAATTACCATCGGAATC 660
DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AGGTGAGCGCGCAGCAATGTCTCATTTAAACCGGTGGGTAAAGCCTCGCATCAATGCTGACG 720

```

```

Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAAGCACAGGAATGCGGAGCGGGTGGGAAAGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGTAAAGTTTGCATAACCGCAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY 841 CGCACGCCAGAGGTTCGCCGCGCAGCCCGCCATCTTATTTGATTTGATTCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCCGCAAAACAGGCGTATGTTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATACTAATCTGGCAATCTCGGCGGCGCCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 AGCTTCCCGGTCCAGCGGATTAACACCGCCAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGCTTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAGCGCGCTGTCTATTAAATACCGCCCGGAGAGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATGTGAAGACGAGGAAATGCGCAGGCGCATGTTGTTGTCGAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCACGCATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 2
Q8GN88 PRELIMINARY; PRT; 432 AA.
ID O8GN88;
AC O8GN88;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AppA.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chen Y., Zhu Z., Zhang Z., He J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF537219; AAN28334.1; -
DR HSSP; P07102; 1DKM.
DR CO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; Hisac_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ SEQUENCE 432 AA; 47055 MW; 5B355D76E7377737 CRC64;

Alignment Scores:

```

```
Pred. No.: 4,88e-172 Length: 432
Score: 2231.00 Matches: 429
Percent Similarity: 99.54% Conservatives: 1
Best Local Similarity: 99.31% Mismatches: 2
Query Match: 94.98% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8GN88 (1-432)

QY 1 ATGAAGCGATCTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCCGCATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGAAAGTGTGTGATTCGATCGCTCATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGCTCAACACGAGCCAGCAACTGATCAGGATGTACCCAGCAGCATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCCCGTAAACCTGGTTGGCTGACACCGCGNGTGTGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTAACCAACGCGAGCTGTGTAGCCGACGGATTGCTGCGGAAAAGGGCTGCCG 300
Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGCTCAGGTCGCGATTATTGCTGATGTGCGAGCGTACCCGTAACAAACAGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGGCTGGACCTGACTGTGCAATACCTATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCGTTATTATCTCTAAACCTGCGCTTTCGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTACGCGATCCTCAGCAGGCGAGGAGGTCAATTCGTGATTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACGCGGTTTCGCGAAGTGGACGCGGTGCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAACACAGGACGAAGCTGTTTCAATACGCGAGCATTAACCATCGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTGACGCGCAATATGCTCAATTAACCGGTGCGTGAAGCTTCGATCAATCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTGCAACACGACGACGAGATGCGGAGCGGGGTGGGAGGATCAC 780
Db 241 LysIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACACGATGGAACCTTCTTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY 841 CGCACGCGCAGGTTGCCGCGCAGCGCCACCCGCTTATTTGATTTGATCGGACGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCGCCATCCACGCAAAACACGCGATCGTGTGATTCACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
```

961 TTTATTGCGGACACGATATACTAATCTGGCAAAATCTGGCGGCGCACTGAGCTCAACTGG 1020
321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
1021 AGCTTCCCGGTACGCGGATTAACAGCGCGGAGTGGTGAAGTGGTGGTGGTGGTGGTGG 1080
341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGlyGluLeuValPheGluArgTrp 360
1081 CGTCGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCCAGACTTTTACAG 1140
361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
1141 CAGATCGGTGATAAAGCGCGTGTCAATTAATACGCCGCCCGGAGAGTGAAACTGACC 1200
381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyGluValLysLeuThr 400
1201 CTGGCAGGATGTCAAGAGCGAATCGCAGGCGATGTTGCTGTCGACGTTTTCAGCAA 1260
401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
1261 ATCGTGAATGAAGCACGATACCGCGTGCAGTTTG 1296
421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 3
Q8RKD6 PRELIMINARY; PRT; 432 AA.
ID Q8RKD6; AC Q8RKD6; DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1] SEQUENCE FROM N.A.
RP MEDLINE=90368616; PubMed=2168385;
RX Dassa J., Marek C., Boquet P.-L.L.;
RA "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
[2] SEQUENCE FROM N.A.
RP MEDLINE=93054596; PubMed=1429631;
RX Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03374; AAA00006.1; -;
DR HSP; P07102; IDNM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc phosphatase.
DR Pfam; PF00328; Acid phosphatase A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolyase.
SQ SEQUENCE 432 AA; 46971 MW; AF86C41EA6193AC5 CRC64;

Alignment Scores:
Pred. No.: 7,09e-172 Length: 432
Score: 2229.00 Matches: 429
Percent Similarity: 99.31% Conservatives: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.89% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8RKD6 (1-432)

QY 1 ATGAACCGATCTTAAATCCCAATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGGCGGAGCTGAAGCTGGAAGTGTGTGATTGTTCAGTCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTCTCAACCAAGGCCACCAACTGATGAGGATGTCCACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACTGGCCGTAATAACCTGGTGTGCTCACCGCGNGTGTGTGAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACCGCAGCTGTGGTAGCCGACCGATGTGGGAAAGGGGTGCGCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTCGTCAAGTCGCGATTATTGCTGATGTCCAGCAGCGTACCCGTAAACACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluAlaThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTCTTAACTCTTAACTGCGGTGGCGAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGAGTCAATTGCTGATTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTCGCGAATCGAACCGGTGCTTAATTTTCGCAATCAAACTTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGAGAAAGCTGTTCATTAAACGAGGATTAACCATCGAACTC 660
Db 201 LeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTAGCCCGCACATGCTCATTAACCGTGCCTGAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGAAATGCGGAGCGGGGTGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACACTTGTGTAAGTTTGCATTAACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCCGAGGTGCGCGCAGCGCCACCCCGTTATTGATTTGATTCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCTATGGTGTGATTCATTCACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATATAATCTGGCAAACTCTCGCGCGCAGCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACCTTCCCGGTACGCGGATACACCGCCGAGGTGGTGAACCTGGTGTGTTGACCGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360

QY 1081 CGTCGCTAAGCCATACAGCCAGTCGATTCAGGTTTCGCTTCCTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATTAACCGCGCTGTCAATTAATACCGCCCGGAGAGGTGAAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGGATGTCAAGACGCGAAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACACGACATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 4
Q8RKD7 PRELIMINARY; PRT; 432 AA.
ID Q8RKD7
AC Q8RKD7
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphonhydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP MEDLINE=93054596; PubMed=1429631;
RA Ostain K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03373; AAA00005.1; -;
DR HSSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SSSEQUENCE 432 AA; 46971 MW; 755D5E4B1AD916A6 CRC64;
Alignment Scores:
Pred. No.: 7,09e-172 Length: 432
Score: 2229.00 Matches: 429
Percent Similarity: 99.31% Conservative: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.89% Indels: 0
DB: 2 Gaps: 0
US-10-034-985-1 (1-1323) x Q8RKD7 (1-432)
QY 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGGCGGAGCTGAAGCTGGAAGTGTGTGATTGTTCAGTCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTCTCAACCAAGGCCACCAACTGATGAGGATGTCCACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACTGGCCGTAATAACCTGGTGTGCTCACCGCGNGTGTGTGAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACCGCAGCTGTGGTAGCCGACCGATGTGGGAAAGGGGTGCGCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTCGTCAAGTCGCGATTATTGCTGATGTCCAGCAGCGTACCCGTAAACACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluAlaThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTCTTAACTCTTAACTGCGGTGGCGAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGAGTCAATTGCTGATTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTCGCGAATCGAACCGGTGCTTAATTTTCGCAATCAAACTTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGAGAAAGCTGTTCATTAAACGAGGATTAACCATCGAACTC 660
Db 201 LeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTAGCCCGCACATGCTCATTAACCGTGCCTGAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGAAATGCGGAGCGGGGTGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACACTTGTGTAAGTTTGCATTAACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCCGAGGTGCGCGCAGCGCCACCCCGTTATTGATTTGATTCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCTATGGTGTGATTCATTCACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATATAATCTGGCAAACTCTCGCGCGCAGCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACCTTCCCGGTACGCGGATACACCGCCGAGGTGGTGAACCTGGTGTGTTGACCGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360

```
QY 121 GTGCGTCTCCAAACGAGCCACGCAACTGATGAGGATGTACCCAGCAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGGTAAACATGGGTGGCTGACACCCCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuAlaTyrLeu 80
QY 241 GGACATTACCAACGCGCTCTGGTAGCCGAGGATTCGTCGCGAAAAGGCTCCCG 300
Db 81 GlyHisTyrGlnAlaGlnArgLeuValAlaAspGlyLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGCTCAGTCGCGATTATTGCTGATGTCGACGAGCGGTACCCGCTAAACAGCGCA 360
Db 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATTAACCTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCGTTATTATTCCTCTAAAACTGGCGGCTTGCCTGCTGATACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTCAGCGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACGCGGTTTCGCAACTGGAACGCGTGTCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTAAACGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGTAGCGCCGACATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTGCAACAGCACAGGAAATCCGAGCGCGGGTGGGGAAGATCAAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTCAACAGTGGACACTTCTTAAGTTGCTAATACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGCACGCGCAGAGTTGCGCGCAGCGCCACCCCGTTATTGATTGATTCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTGATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTATTCCGCGACGATACATACTCGCAATCTCGCGCGCGCAGCTGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGCTCAGCCGAGTAACACCGCGCAGGTGGTGAACCTGGTGTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGTAAACGATAAACAGCACTGATTCAGTTTCGCTGGCTTCCTCCAGCTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGGTGATAAAGCCGCTCTCATTAATACCGCGCGCGGAGAGGTCAAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
```

```
QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGGGATGTGTTCGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATCAAGCAGCATACCGCGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 5
Q8RKD8 PRELIMINARY; PRT; 432 AA.
ID Q8RKD8
AC Q8RKD8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostainin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL; L03372; AAA00004.1; -.
DR HSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phspstse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
DR Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; 9A85536B57FCCFB5 CRC64;
Alignment Scores:
Pred. No.: 7,09e-172 Length: 432
Score: 2229.00 Matches: 429
Percent Similarity: 99.31% Conservative: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.89% Indels: 0
DB: 2 Gaps: 0
US-10-034-985-1 (1-1323) x Q8RKD8 (1-432)
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATGTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTCTCAACCAAGCCAGCCAGCATGATCAGAGTGTCAACCCAGCAGCATGGCCA 180
Db 41 ValAlaAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCGGTAAACATGGGTGGCTGACACCCCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuAlaTyrLeu 80
```

```
QY 241 GGACATTACCAACGCCAGCGCTGTTAGCCGACGCGATTGCTGGCGAAGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnA-gLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCCAGTCCGCGATTATTGCTGATGTCACGAGCGTACCCGTAACAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGCTGGCAGCTGCTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATCTCTAAACTCGCGTTCGCCAATCGATTAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTACGCGATCTCTCAGCAGCGCAGGAGGTCATTTGCTGATTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTCGCGAATCGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCTATTACGAGGATTAACCATCGAACTC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCCCGCACATGCTCATTAACCGTGCCTGAGCTCGCATCAATCGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTCGCAACAGCACAGGGAATCGCGAGCGCGGTGGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACGAGTGGAACACTTGCTAAGTTTGATTAACCGCGCAATTTTATTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCCGAGGTGCGCCGAGCGCGCCACCCGTTGATTTGATTTGATGATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCTATGCTGACATTAACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTATTGCGGACACGATATAATCTCGCAAAATCTCGCGCGCGCAGCTGAGCTCACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACCTTCCCGTCAGCCGGATACACGCCCGAGGTGTTGAACTGGTGTGTTGACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCAGTAAACAGCCAGTTCAGTTTCGTTGCTGCTTCCAGACTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGCGTGATAAAACCGCGCTGTCTAATAACGCCCGCGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGCGGAATGCGCAGCGCATGTTGCTGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATACGACCATACCGCGCTGCGATTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 6

```
Q8RKE0
ID Q8RKE0 PRELIMINARY; PRT; 432 AA.
AC Q8RKE0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP MEDLINE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP MEDLINE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostain K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03370; AAA00002.1; -.
DR HSP; P07102; IDKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_pheptse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; 5BB632D4682EABF CRC64;
```

Alignment Scores:
Pred. No.: 7,09e-172 Length: 432
Score: 2229.00 Matches: 429
Percent Similarity: 99.31% Conservative: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.89% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8RKE0 (1-432)

```
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGGTGTGTCAGTCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerAlaHisGly 40
QY 121 GTCGCTGCTCCAAACCAAGCCAGCAACTGATGAGGATCTACCCCGAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTGGCTGCACACCGCGNGGTGGTGAAGTAACTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTACCAACCGCAGCGTCTGGTAGCCGAGGATGCTGGCGGAAAAGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCCAGTCCGCGATTATTGCTGATGTCGAGCGGTACCCGTAACAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
```

```

Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTAATCTCTAAAACTGGCGTTTGGCAACTGGTAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGACTACCGGATCTCCACAGGCGAGGAGGTCAATTGCTGACTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerAgaIaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGGCGTTTCGGCAACTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTACGAGGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGACATGCTCATTAACCGGTGGGTAGCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGTGGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGCTAAGTTTGCATTAACCGCAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGCACGCGCAGAGTTGCGCGCAGCGCCACCCGCTTATTTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCGCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCGCGCACGATCTATCTGCGAAATCTCGCGCGCGCAGTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGTCAGCGGATAACACGCGCCAGTGGTCAACTGGTGTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCCGCTAAGCGATAACACCCAGTGGATTTCAGTTTCGTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGCTGATAAAACCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGSCAGGATGTAAGAGCGAAATACGCGAGGATGTTGCTGGCAGGTTTATACCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCACCATACCGCGCTGCGTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

```

RESULT 7

```

Q8RKD9 PRELIMINARY; PRT; 432 AA.
AC Q8RKD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
RN NCBI_TaxID=562;
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03371; AAA00003.1; -.
DR HSSP; P07102; 1DKO.
DR GO; GO:0003993; P:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phspstse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE. 432 AA; 47033 MW; 9F29B9DF9C368175 CRC64;

```

```

Alignment Scores:
Pred. No.: 8,55e-172 Length: 432
Score: 2228.00 Matches: 429
Percent Similarity: 99.54% Conservative: 1
Best Local Similarity: 99.31% Mismatches: 2
Query Match: 94.85% Indels: 0
DB: 2 Gaps: 0

```

US-10-034-985-1 (1-1323) x Q8RKD9 (1-432)

```

QY 1 ATCAAGAGCATCTTAATCCCATTTTATCTTCTGATTCCGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTGCTGCTGCTATGCT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgAsnGly 40
QY 121 GTGCTGCTCCAAACCAAGCCACGCAACTGATCGCAGATGTCAACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACTGGGTTCGTGACACCGCGNGGTGCTGAGCTTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACCCAGCGCTCTGGTAGCCGACGAGTTGTCGCGAAAAGGGCTGCCCG 300
Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
QY 301 CAGTCTGCTCAGTCCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGGCACCTGCTGCTGCAATAACCTACATACCCAGGACAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGAGCGGCTTTCGCAACTGGATAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGACTACCGGATCTCCACAGGCGAGGAGGTCAATTGCTGACTTACGGGCAT 540

```

```
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACGGCGTTCCGGAACCTGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAACAGGACGAAAGCTGTTTCATTAACGCGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCGCAATGCTCTAATTAACCGTCCGCTGAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaPheAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTGCAACAGCACAGGAATGCCGAGCGCGGGTGGGAAGATCAACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Qy 781 GATTACACAGTGGACACCTTGCTAAGTTTGCATTAACCGGCAATTTATTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Qy 841 CGCAGCGCAGAGGTGCGCGCAGCGCCACCCGTTATTGGATTGATGATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Qy 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGCTGATACATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTCCCGCAGCATACTAATCTGGCAAACTCGCGCGGCGACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACCTTCCCGTCCAGCGGATACACCGCCAGTGGTGAACCTGGTGTGAACTGGTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGTAAGCATACAGCCAGTGTGATTCAGTTTCGTTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATCGGTGATAAAGCGCGTGTCTAATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGAGGATGTGAGACGGAATGCCAGGCGATGTTGCTGTCGGCAGGTTTACGCA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAAGCAGCATACCGCGCTGAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 8

```
Q8RKD5 PRELIMINARY; PRT; 432 AA.
ID Q8RKD5 AC Q8RKD5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP MEDLINE=9038616; PubMed=2168385;
RX Dassa J., Marck C., Boquet P.-L.B.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
```

```
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
```

```
DR EMBL: L03375; AAA00007.1; -.
DR HSP; P07102; 1DKM.
DR GO: 0003993; F: acid phosphatase activity; IEA.
DR GO: 0016787; F: hydrolase activity; IEA.
DR InterPro: IPR000560; HisAc_phsphtse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46990 MW; 951F393EA9A1A47C CRC64;
```

Alignment Scores:

```
Pred. No.: 1-5e-171 Length: 432
Score: 2225.00 Matches: 429
Percent Similarity: 99.31% Conservative: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.72% Indels: 0
DB: 2 Caps: 0
```

US-10-034-985-1 (1-1323) x Q8RKD5 (1-432)

```
Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGGTGTGTCAGTCTGCTCAGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTCGGTGCTCAACCAAGGCCAGCAACTGATGAGGATGTACCCCGAGCGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGCGCGGTAAACCTGGGTTGGCTGACACCCCGNGGTGGTGAAGTAACTGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Qy 241 GGACATTACCAACCGCAGCTCTGCTAGCCGAGGATGCTGGCGAAAGGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTCAGCAGCGGTACCCGTAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGGCTGGCACCCTGATGTGCAATAACCGTACATACCCAGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCAGTCCCGATCCGTTATTATCTTAATAAAGTGGCGGTTTGGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTGACCGGATCTCTCAGCAGGCGAGGAGGTCAATTCTGCTGCTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGATTAACGAGGATTAAC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```


QY 661 AAGTGAGCCGACCAATGCTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGAAATCCGAGGCGCGGTGGGGAAGATGACG 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACACAGTGGACACTGCTGAAGTTTCATTAACCGCGCAATTTATTTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGACGCGCACAGGTGGCCCGCAGCCGCGCACCCCGTTATTGGATTGATCATGGCAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCCCATCCACCGCAAAACAGCGCTATGGTGTGATCATCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGACGACATATACTGGCAATCTCGCGCGCCTGAGGCTCACTGG 1020
 DB 321 PheIleAlaGlyAlaAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTCTCCCGGTGAGCGGATAACACGCGCCAGGTGGTGAACCTGGTGTGAACTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGTAAGCGGATAACACCGAGTGGATTGAGTTTCGTGCTTCCAGACTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGGTGTATAAACCGCGCTGTCTATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGAGGATGTGAAGAGCGAATCGCAGGGATGTTTCGTGGCAGGTTTACGCAA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAACGACCATACCGCGTGCAGTTTG 1296
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 RESULT 9
 Q83RW2 PRELIMINARY; PRT; 442 AA.
 AC Q83RW2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name:app; OrderedLocusNames=SF0982;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441 (2002).
 DR EMBL; AE015127; AAN42610.1; -;
 DR HSSP; P07102; 1DKL.

DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR InterPro; IPR000560; HisAc_phosphatase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW complete proteome.
 SQ SEQUENCE 442 AA; 48183 MW; 20BCE2C454254680 CRC64;
 Alignment Scores:
 Pred. No.: 2, 63e-171 Length: 442
 Score: 2222.00 Matches: 427
 Percent Similarity: 99.07% Conservative: 1
 Best Local Similarity: 98.84% Mismatches: 4
 Query Match: 94.59% Indels: 0
 DB: 2 Gaps: 0
 US-10-034-985-1 (1-1323) x Q83RW2 (1-442)
 QY 1 ATGAAGCGCATTAATCCCATTTTATCTCTTCTGATTCCTGTTAAACCCGCAATCTGCA 60
 DB 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
 QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCAGTCCGTCATGCT 120
 DB 31 PheAlaGlnSerGluProGluLeuLysLeuGluSerMetValIleValSerArgHisGly 50
 QY 121 GTGCGTGTCTCAACACGAGCCACGCAATCATGATCAGAGATGTCAACCCAGACGATGGCCA 180
 DB 51 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 70
 QY 181 ACCTGCGCGTAAACATGGGTTGGCTGACACCGCGNGGTGGTGAAGTAAATCGGCTATCTC 240
 DB 71 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 90
 QY 241 GGCATTAACCAACCCAGCGTCTGGTAGCCGAGATTGCTGCGGAAAAGGGCTGCCCG 300
 DB 91 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 110
 QY 301 CAGTCTGGTCCAGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAACAAACAGCGCAA 360
 DB 111 GlnSerAlaGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 130
 QY 361 GCCTTCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTATACCCAGCGAGATACG 420
 DB 131 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 150
 QY 421 TCCAGTCCCGATCCGTTATTATTAATCCTCTAAAACCTGGCGTTCGCAACTGGATAACGCG 480
 DB 151 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 170
 QY 481 AACGTGACTGACGCGATCCCTCAGCAGGCGCAGGAGGTCAATTTGCTGACTTTACCGGCAT 540
 DB 171 AsnValThrAspAlaIleLeuCysArgAlaGlyGlySerIleAlaAspPheThrGlyHis 190
 QY 541 CGGCAAAACCGGTTTCGCGAATCGGAACGCGGTGCTTAATTTTCGCGAATCAAACTGTGCG 600
 DB 191 ArgGlnThrValPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 210
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTACGCGAGCATTAACCATCGGAATC 660
 DB 211 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 230
 QY 661 AAGGTAGCGCCGACATGCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATCTGACG 720
 DB 231 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 250
 QY 721 GAGATATTTCTCTGCAACAGCACAGGAAATCCGAGCGCGGTGGGAGGATCACC 780
 DB 251 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 270
 QY 781 GATTACACAGTGGAAACACTGCTTAAGTTTGCATAACCGGCAATTTTATTTGCTACAA 840
 DB 271 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 290

QY 841 CGCAGCCAGAGGTTGCCCGAGCCGCCCAACCCCGTTATTGGATTGATCGGCAGCG 900
 DB 291 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 310
 QY 901 TTGAGCCGCCATCCACCGCAAAACACGCGGTGATGGTGACATTACCCACTTCAGTACTG 960
 DB 311 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 330
 QY 961 TTTATTCCCGACACGATACTAATCTCGCAAACTCTCGCGCGCGCACTGGAGCTCAACTGG 1020
 DB 331 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 350
 QY 1021 ACGTTCCCGGTGAGCCGAGTAAACACCGCCAGGTGGTGAACCTGGTGTGAAACGCTGG 1080
 DB 351 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 370
 QY 1081 CGTCGGCTACGATACACGCCAGTGGATTTCAGGTTTCGGTCTTCAGACTTTACAG 1140
 DB 371 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 390
 QY 1141 CAGATGGCTGATAAAACCGCGCTGTCTAATAACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 391 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 410
 QY 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTGTTGTTGCGAGGTTTACGCAA 1260
 DB 411 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 430
 QY 1261 ATGCTGAATCAGCACGCATACCGGCTGCAGTTTG 1296
 DB 431 IleValAsnGluAlaArgIleProAlaCysSerLeu 442

RESULT 10

Q6RK08 PRELIMINARY; PRT; 432 AA.
 ID Q6RK08 AC Q6RK08;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Acid phosphatase/phytase 2.
 GN Name=appa2;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99194564; PubMed=10092520; DOI=10.1006/bbrc.1999.0361;
 RA Rodriguez E., Han Y., Lei X.G.;
 RT "Cloning, sequencing, and expression of an Escherichia coli acid
 RT phosphatase/phytase gene (appa2) isolated from pig colon.";
 RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lei X.G.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY496073; AAR87658.1; -
 DR GO: GO:0003993; P:acid phosphatase activity; IEA.
 DR InterPro: IPR005060; HisAc phosphatase.
 DR Pfam: PF00328; Acid phosphatase A; 1.
 DR PROSITE: PS00616; HIS ACID PHOSPHAT 1; 1.
 DR PROSITE: PS00778; HIS ACID PHOSPHAT 2; UNKNOWN 1.
 FT CHAIN 20 432
 SQ SEQUENCE 432 AA; 47041 MW; 71B0E5ED2EA2674 CRC64;
 Alignment Scores:
 Pred. No.: 3.16e-171 Length: 432
 Score: 2221.00 Matches: 428
 Percent Similarity: 99.07% Conservativeness: 0
 Best Local Similarity: 99.07% Mismatches: 4
 Query Match: 94.55% Indels: 0
 DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q6RK08 (1-432)
 QY 1 ATCAAGAGCATCTTAATCCCATTTTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCAGTCGTCATGGT 120
 DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTCGCTGCTCCACCAAGCCAGCCAGCACTGATCGAGGATGTCAACCCAGAGCATGGCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCGCGGTAAACCTGGGTTGCTGACACCCCGNGGTGCTGAGCTAACTCGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGACATTTACCAACCGCAGCGTCTGGTAGCCGCGATTGCTGCGGAAAAAGGGCTGCCCG 300
 DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGCTCAGTCTCGGATTATTGCTGATGTCGACGAGGTACCCGTAAACAGCGCAA 360
 DB 101 GlnProGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGCGCACCTGCTGCAATAACCGTACATACCCAGGCGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCCGCATCCGTTATTATCTCTAAACAACTGGCGTTTCCCAACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
 QY 481 AACGTGACTGACCGCATCTCTACGAGGCGAGGAGGTCAATTCTGACTTTTACGGGCAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGCAACACCGCGTTTCGCAACTGGAACCGGTCTTAATTTTCCGCAATCAAACTGTGTC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
 QY 601 CTTAAACGTCGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGATTTACCATCGGAATC 660
 DB 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AGGTGAGCGCGCAATGCTCTCAATTAAACGCTGCGGTAAAGCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGGAAGGATCACCC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACACGAGTGAACACTTGTCTAAGTTTTCATAACGCGCAATTTTATTTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 QY 841 CGCACGCCAGAGTTCCCGCGAGCCCGCGCACCCCGTTATTGATTTGATCGGAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACCCCGCATCCACCGGAAAAACAGGCGTATGTTGTGACATTCCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACACGATACTAATCTGGCAAACTCTCGCGCGCGCACTGGAGCTCAACTGG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTCCCGGTGAGCGGATACACCGCGCGAGGTGGTGAACCTGGTGTGAAACGCTGG 1080

```

Db      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY      1081 CGTGGTAAAGCATAACACCGAGTGGATTTCAGGTTTCGGTGTCTTCCAGACTTTACAG 1140
Db      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY      1141 CAGATCGGTGATAAACCGCGCTGTTCATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
Db      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY      1201 CTGGCAGGATCTCAGAGCCAAATCGCAGGCGATGTCGTTCGCGAGGTTTACGCA 1260
Db      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY      1261 ATCGTGAATCAGCAGCATACCGCGTGCAGTTTG 1296
Db      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 11
Q7UD08
ID      Q7UD08      PRELIMINARY;      PRT;      432 AA.
AC      Q7UD08
DT      01-OCT-2003 (TrEMBLrel. 25, Created)
DT      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
GN      Name=appa; OrderedLocusNames=S1048;
OS      Shigella flexneri.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Shigella.
OX      NCBI_TaxID=623;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=2457T;
RX      MEDLINE=22590274; PubMed=12704152;
RX      DOI=10.1128/JAI.71.5.2775-2786.2003;
RA      Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA      Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA      Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA      Schwartz D.C., Blattner F.R.;
RA      "Complete genome sequence and comparative genomics of Shigella
RT      flexneri serotype 2a strain 2457T.";
RL      Infect. Immun. 71:2775-2786(2003).
DR      EMBL; AE016981; AAP16495.1; -.
DR      HSP; P07102; 1DKL.
DR      GO; GO:0003993; F-actin phosphatase activity; IEA.
DR      InterPro; IPR000560; HisAc_phsphtse.
DR      Pfam; PF00328; Acid_phosphat_A; 1.
DR      PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
DR      PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ      SEQUENCE 432 AA; 47136 MW; 20042239BC66DFC CRC64;

Alignment Scores:
Pred. No.:      5.53e-171      Length:      432
Score:          2218.00      Matches:      426
Percent Similarity: 98.84%      Conservative: 1
Best Local Similarity: 98.61%      Mismatches: 5
Query Match:     94.42%      Indels:      0
DB:              2      Gaps:        0

US-10-034-985-1 (1-1323) x Q7UD08 (1-432)
QY      1 ATGAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGGTTAAACCCCGCAATCGCA 60
Db      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY      61 TTCCTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGTGATTGTCAGTCGTCTGATG 120
Db      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerMetValIleValSerArgHisGly 40
QY      121 GTGGCTCTCCACACAGGCCACCACTGATGTCAGGATGTCAACCCAGACGCATGCCA 180
Db      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrp 60

```

```

QY      181 ACTGCGCCGGTAAACCTGGGTGGCTGACACCCCGNGGTGGTGGAGTAAATCGCCTATCTC 240
Db      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY      241 GGACATTACACGCCAGCGCTGTGTAGCCGCGGATTCTGCGGAAAAGGCGTCCCG 300
Db      81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY      301 CAGTCTGGTCAGTCCGCGATTATTGCTGATGTCGACGAGCGTACCGCTAAAAACAGCGAA 360
Db      101 GlnSerAlaGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY      361 GCCTTCGCCCGCGGCTGCGACCTGACTGTGCAATAACCGTACATACCAGGCGAGATACG 420
Db      121 AlaPheThrAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY      421 TCCAGTCCCGGATCCGTTATTAACTCTTAAAAACTGGCGTTTGCACACTGGATAACGG 480
Db      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY      481 AACGTGACTGACGCGATCTCAGACGGCAGGAGGTCAATTGCTGACTTTACCGGCGAT 540
Db      161 AsnValThrAspAlaIleLeuCysArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY      541 CGGCAAAACCGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTGTGC 600
Db      181 ArgGlnThrValPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY      601 CTTAAACGTGAGAAACAGCAGCAAGAACTGTTCAATTAACCGAGCATTTACCATCGGAATC 660
Db      201 LeuAsnArgLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY      661 AAGGTGAGCGCGCACAAATGCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATGTCGACG 720
Db      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY      721 GAGATATTTCTCTGCAACAGCAGCAAGGAATGCCGAGCGCGGGTGGGGAAGATCACC 780
Db      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY      781 GATTCAACACGAGTGAACACCTTGCTAAGTTTGCATTAACCGCAATTTTATTGCTACAA 840
Db      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY      841 CGCACGCCAGAGGTTGCCCGCAGCCGCGCACCCCGTTATTGATTTGATTCATGCGACGG 900
Db      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY      901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCACCTTCAGTACTG 960
Db      301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
QY      961 TTTATTTCGGGACACGATCAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Db      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
QY      1021 ACGCTTTCGGGTGATAAACCGCGCTGCTAATAATACCGCGCGAGAGGTGAAACTGACC 1080
Db      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY      1081 CGTGGCTAAGCGAATAACGCCAGTGGATTTCAGGTTTCGGTGTCTTCAGACTTTACAG 1140
Db      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY      1141 CAGATCGGTGATAAACCGCGCTGCTAATAATACCGCGCGAGAGGTGAAACTGACC 1200
Db      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY      1201 CTGGCAGGATGTGAAGACGCGAAAATGCGAGGCGATGTTGCTGTGGCAGGTTTTACGCCAA 1260
Db      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

```

```
QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTG 1296
Db 421 ILeValAenGluAlaArgIleProAlaCysSerLeu 432
RESULT 12
ID Q8CW75 PRELIMINARY; PRT; 446 AA.
AC Q8CW75;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic appA protein.
GN Name=appA; OrderedLocuNames=cl121;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016758; AAM79589.1; -.
DR HSP; P07102; 1DKQ
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc phosphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 446 AA; 48605 MW; F1308CD691DB9F49 CRC64;

Alignment Scores:
Pred. No.: 1.42e-170 Length: 446
Score: 2213.00 Matches: 425
Percent Similarity: 99.07% Conservative: 3
Best Local Similarity: 98.38% Mismatches: 4
Query Match: 94.21% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8CW75 (1-446)
QY 1 ATGAAGCGATCTTAATCCCAATTTTATCTTCTGATTCGTTAAACCCGCAATCGCA 60
Db 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProLysSerAla 30
QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCAGTCTCATGGT 120
Db 31 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 50
QY 121 GTGCGTGTCTCAACCAAGGCCACCACTGATGACGAGTGTACCCACGACGATGGCCA 180
Db 51 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 70
QY 181 ACCTGGCCGGTAAACATGGTGTGCTGACACCGCGNGGTGGTGAAGTAACTCGCTATCTC 240
Db 71 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 90
QY 241 GGACATTACCAACCCGCGCTGCTGATACCGACGATTCGCGCAAAAGGCTGCGCG 300
Db 91 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGlyCysPro 110
QY 301 CAGTCTGGTCAAGTCCGGATATTGCTGATGTCACGAGCGGTACCCGTAACCAAGGCGAA 360
Db 111 GlnProGlyGlnValAlaIleIleSerAspValAspGluArgThrArgLysThrGlyGlu 130
QY 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGTGCAATAACCGTACATACCCAGCAGATACG 420

131 AlapheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 150
421 TCCAGTCCCGATCCGTTTATTAATCTCTTAAAAAAGCTGGCGTTTGGCCAACTGGATAAGCGC 480
151 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 170
481 AACGTGACTGACGGATCTCTCAGCAGGCGCAGGAGGTCAATTCGTGACTTTTACCGGGCAT 540
171 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 190
541 CGGCAACCGCGTTTCGCGAACCTGGACCGGTCTTAATTTCCGCAATCAAACTTCGTC 600
191 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 210
601 CTTTAAACGCTGAGAAACAGGACGAAAGCTGTTCATTAAACGAGCGCATTAACCATCGAACTC 660
211 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 230
661 AAGTGAGCCCGCAACATGCTCTATTAAACCGTGGGTAAAGCTTCGCATCAATGCTGACG 720
231 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 250
721 GAGATATTTCTCTGCAACACAGCACAGGGAATCCGCGCGCGGCTGGGGAAGCATCAC 780
251 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 270
781 GATTTCACACAGTGAACACCTTGTAAAGTTTGCATAACCGCAATTTTATTTGCTACAA 840
271 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 290
841 CGCACCGCAGAGTTGCCCGCACCGCCACCCGCTATTGATTTGATTCATGCGCAGCG 900
291 ArgThrProGluValAlaArgSerAlaThrProLeuLeuAspLeuIleMetAlaAla 310
901 TTGACGCCCATCCACCGCAAAACAGCGGTATGCTGTGACATTACCCTTCAGTACTG 960
311 LeuThrProHisProSerGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 330
961 TTTTATTCGCGCACGATCTAATCTGGCAAAATCTCGCGCGCGCACCTGGAGCTCAACTGG 1020
331 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 350
1021 ACGTCTCCCGTCCAGCGGATAACACGCGCCAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
351 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 370
1081 CGTCGCTAAGCGATACACGACGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
371 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 390
1141 CAGATCGTGTATAAACCGCGCTGTCTAATAATAGCGCGCGGAGAGGTGAACACTGACC 1200
391 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 410
1201 CTGCGCAGGATGTGAAGACGCGAAATGCGCAGGCGATGTGTTCTGTTGGCAGGTTTTACGCA 1260
411 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 430
1261 ATCGTGAATGAACGACGATACCGCGCTGCGAGTTG 1296
431 IleValAsnGluAlaArgIleProAlaCysAlaLeu 442

RESULT 13
ID Q7AFW3 PRELIMINARY; PRT; 434 AA.
AC Q7AFW3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phosphoanhydride phosphorylase.
GN OrderedLocuNames=ECs1136;
OS Escherichia coli O157:H7.
```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Havaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AP002554; BAB34559.1; -;
 DR GO: GO:0003993; P:acid phosphatase activity; IEA.
 DR InterPro: IPR000560; HisAc_phosphatase.
 DR Pfam: PF00328; Acid phosphatase A; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
 SQ SEQUENCE 434 AA; 47337 MW; F197DF7D1869F9C4 CRC64;

Alignment Scores:
 Pred. No.: 2,46e-170 Length: 434
 Score: 2210.00 Matches: 427
 Percent Similarity: 98.39% Conservative: 0
 Best Local Similarity: 98.39% Mismatches: 5
 Query Match: 94.08% Indels: 2
 DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q7AFW3 (1-434)

QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCGCAATCTGCA 60
 DB 1 MetLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGTGATTGTGAGTCGT 114
 DB 21 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 40
 QY 115 CATGGTGTGGTGTCAACCAAGGCCAGCACTGATGAGAGTGTACCCCGAGCGCA 174
 DB 41 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 60
 QY 175 TGCCCAACCTGGCCGTTAAACCTGGTGTGACCGCGNGTGTGTGAGCTAATCGCC 234
 DB 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 80
 QY 235 TATCTCGGACATTAACCAACGCGCTCTGTAGCCGAGGATTCGTGGCGAAAAGGCG 294
 DB 81 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 100
 QY 295 TGCCCGGAGTGTGTGAGTGTGCGATTTGCTGATGTGACGAGCGTACCGGTAAACA 354
 DB 101 CysProGlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThr 120
 QY 355 GGCGAAGCTTCGCGCGCGCTGCGACCTGACTGTGCAATTAACCGTACATACCCAGCA 414
 DB 121 GlyIleAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 140
 QY 415 GATACGTCGATCCGATCCGTTATTTAATCTCTAAAACTGGCGTTGCGCACTGGAT 474
 DB 141 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 160
 QY 475 AACCGGACGTGACTGACGCGATCCTGACGAGGCGAGGAGGTCAATTGCTGACTTTACC 534
 DB 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 180
 QY 535 GGGCATCGCAACCGGGTTTCGGGAACCTGGAACGGGTGCTTAATTTTCGCAATCAAC 594
 DB 181 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 200
 QY 595 TTGTGCTTAAACGTGAGAAACAGGACGAAGAGCTGTTTCATTAAACGAGGCATTACCATCG 654

DB 201 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 220
 QY 655 GAACCTCAAGGTGAGCGCGCACCAATGCTCATTAACCGGTGCGTAAGCCTCGCATCAATG 714
 DB 221 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
 QY 715 CTGACCGAGATATTTCTCTCGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGG 774
 DB 241 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 260
 QY 775 ATACCGATTACACCGAGTGAACACTGCTTAAGTTTGCATTAACCGGCAATTTATTCG 834
 DB 261 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeu 280
 QY 835 CTACACGACGCGCAGAGGTTGGCCGAGCGCGCCACCCCGCTATTGATTTGATCATG 894
 DB 281 LeuGlnArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 300
 QY 895 GCAGCGTTGACCGCCCATCCACCGCAAAACAGGCGTATGTTGTGACATTACCACCTCA 954
 DB 301 IleAlaLeuThrProHisProGlnLysGlnAlaThrGlyValThrLeuProThrSer 320
 QY 955 GTRACTGTTATTCGCGACACGATACTACTGCGCAATCTCGCGCGCGACTGGAGCTC 1014
 DB 321 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 340
 QY 1015 AACTGGACGCTTCCCGGTGACCGGATAACACCGCCCGAGTGTGTGAACCTGGTGTGAA 1074
 DB 341 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 360
 QY 1075 CGCTGGCGTGGTAAAGCGATTAACACGAGTGGATTCAGGTTTCGTGTTCCAGACT 1134
 DB 361 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 380
 QY 1135 TTAACGAGATGCTGATATAAACCGCGCTGTCTATTAAATACGCGCGCGAGAGGTGAA 1194
 DB 381 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 400
 QY 1195 CTGACCTGCGCAGGATGTGAAGAGCGAAATGCCAGCGCATGTGTTCGTTGGCAGGTTT 1254
 DB 401 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 420
 QY 1255 ACGCAATCTGTAATGAAGCACCGCATACCGCGTGCAGTTTG 1296
 DB 421 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 434

RESULT 14
 Q8XC29 PRELIMINARY; PRT; 444 AA.
 AC Q8XC29;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphatase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name:appA; OrderedLocusNames=z1397;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RL "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).

DR EMBL; AE005292; AAC55528.1; --
 DR PIR; D85633; H90770.
 DR PIR; H90770; H90770.
 DR HSSP; P07102; 1DKL.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR Pfam; PF00328; Acid phosphat A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 444 AA; 48388 MW; A9AA7E6653AP247E CRC64;

Alignment Scores:
 Pred. No.: 2,48e-170 Length: 444
 Score: 2210.00 Matches: 427
 Percent Similarity: 98.39% Conservative: 0
 Best Local Similarity: 98.39% Mismatches: 5
 Query Match: 94.08% Indels: 2
 DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q8XC29 (1-444)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
 |||||
 DB 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
 |||||

QY 61 TTGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTCGT 114
 |||||
 DB 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 50
 |||||

QY 115 CATGGTGTGCTGCTCAACCAAGGCCACCACTGATGAGGATGTCACCCAGAGGCA 174
 |||||
 DB 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
 |||||

QY 175 TGGCAACCTTGGCCGGTAAACCTGGGTGGCTGACACCGCGGTGTGAGTAACTCGCC 234
 |||||
 DB 71 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 90
 |||||

QY 235 TATCTCGACATTACCAACCCGCGCTGTGATGCGCAGCGATGTCGGCGAAAGGCG 294
 |||||
 DB 91 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 110
 |||||

QY 295 TGCCCGCAGTCTGGTCAGTCTGGGATTTGCTGATGTCACGAGGTACCCGTAACAA 354
 |||||
 DB 111 CysProGlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThr 130
 |||||

QY 355 GCGAAGCCTTCCGCGCGGCTGGCACCTGCTGCAATACCGTACATACCCAGGCA 414
 |||||
 DB 131 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 150
 |||||

QY 415 GATACGTCCAGTCCCGATCCGTTATTTAATCCTCTAAATACTGGCGTTTCCCACTGGAT 474
 |||||
 DB 151 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 170
 |||||

QY 475 AACCGAAGTGTACTGACGATCTCCTCAGAGGCGAGGAGGTCAATGCTGATCTTACC 534
 |||||
 DB 171 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 190
 |||||

QY 535 GGCATCCGCAACCGGTTTCGCAACTGCAACGGGTGCTTAATTTCCGCAATCAAC 594
 |||||
 DB 191 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 210
 |||||

QY 595 TTGTCCTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTAACGAGGATTAACCATCG 654
 |||||
 DB 211 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 230
 |||||

QY 655 GAATCAAGTGTAGCCCGCAATGTCTCAATTAACGGTGGGTAGCTCGCATCAATG 714
 |||||
 DB 231 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 250
 |||||

QY 715 CTGACCGAGATATTCTCTCGACACGACGAGGATCGGAGCGGGTGGGAGG 774
 |||||
 DB 251 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 270
 |||||

775 ATCACCGATTACACACAGTGGAAACACCTTGCTAAGTTTGTGATTAACCGGCAATTTTATTG 834
 |||||
 DB 271 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 290
 |||||

QY 835 CTAAACGACGACGACGAGGTGCCCCGACGCGCCACCCGTTATTGGATTGATCATG 894
 |||||
 DB 291 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 310
 |||||

QY 895 GCAGCGTGTACGCGCCCATCCACCGCAAAACACGCGGTATGTTGTGACATTACCCACTTCA 954
 |||||
 DB 311 IleAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 330
 |||||

QY 955 GTRACTGTTATTCGCGACACGATACTAATCTGCGAAATCTCGCGCGGCACTGGAGCTC 1014
 |||||
 DB 331 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 350
 |||||

QY 1015 AACTGACGCTTCCCGTCCGCGGATACACACCGCCGAGTGTGTAACCTGGTGTGAA 1074
 |||||
 DB 351 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 370
 |||||

QY 1075 CGCTGCGTGGCTAAAGCGATACACGCGAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACT 1134
 |||||
 DB 371 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 390
 |||||

QY 1135 TTACACGATGCGTGTATAAACGCGCTGTCTAATTAATACGCGCCGCGGAGAGGTGAAA 1194
 |||||
 DB 391 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 410
 |||||

QY 1195 CTGACCTCGCAGGATGTGAAGCGAATACGCGAGGATGTCGCGGATGTCGTTGGCAGGTTT 1254
 |||||
 DB 411 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 430
 |||||

QY 1255 ACCCAATCTGTAATGAAGCACGATACCGCGCTGCAGTTTG 1296
 |||||
 DB 431 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 444
 |||||

RESULT 15
 Q676V7 PRELIMINARY; PRT; 433 AA.
 ID Q676V7
 AC Q676V7;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Phytase.
 GN Name=PhyA;
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VKPM B-4090;
 RA Zinin N.V., Syneoki S.P.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY390262; AAR89622.1; -
 DR InterPro; IPR000560; HisAc_phsphtse.
 DR Pfam; PF00328; Acid phosphat A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 SQ SEQUENCE 433 AA; 48506 MW; A7923288F2PCFD44 CRC64;

Alignment Scores:
 Pred. No.: 1,32e-102 Length: 433
 Score: 1374.50 Matches: 265
 Percent Similarity: 73.15% Conservative: 51
 Best Local Similarity: 61.34% Mismatches: 115
 Query Match: 58.51% Indels: 1
 DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q676V7 (1-433)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
 |||||
 |||||

```
Db      1 MetSerThrPheIleIleArgLeuLeuPhePheSerLeuLeuCysGlySerPheSerIle 20
Qy      61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGGTGAATGTCAGTCTCATGTT 120
Db      21 HisAlaGluGlnAsnGlyMetLysLeuGluArgValValIleValSerArgHisGly 40
Qy      121 GTGGGTGCTCAACCAAGGCCAGCACTGATCAGGATGTCAACCCAGCAGCATGGCCA 180
Db      41 ValArgAlaProThrLysPheThrProIleMetLysAspValThrProAspGlnTrpPro 60
Qy      181 ACCTGGCCGTAAACATGGGTGGCTGACACCCGCGNGTGGTGAAGTAAATCGCTATCTC 240
Db      61 GlnTrpAspValProLeuGlyTrpLeuThrProArgGlyGlyGluLeuValSerGluLeu 80
Qy      241 GGCATTTACCAAGCCAGCTCTGGTAGCCGAGGATGCTGGCGAAAGAGGCTGCCCG 300
Db      81 GlyGlnTrpGlnArgLeuTrpPheThrSerLysGlyLeuLeuAsnAsnGlnThrCysPro 100
Qy      301 CAGTCTGGTCAGGTTCGGGATTATTGCTGATGTCGACGAGGTACCCGCTAAACAGGCGAA 360
Db      101 SerProGlyGlnValAlaValIleAlaAspThrAspGlnArgThrArgLysThrGlyGlu 120
Qy      361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATACCGTACATACCCAGCAGCATACG 420
Db      121 AlaPheLeuAlaGlyLeuAlaProLysCysGlnIleGlnValHisTyrGlnLysAspGlu 140
Qy      421 TCCAGTCCCGATCGTGTATTATTAATCTCTAAACATGGCGTTCGCCAACTGGATACGGC 480
Db      141 GluLysThrAspProLeuPheAsnProValLysMetGlyThrCysSerPheAsnThrLeu 160
Qy      481 AAGCTGATCAGCGATCTCTCAGCAGGCGAGGAGGTCAATGTGCTACTTTTACCGGCGAT 540
Db      161 LysValLysAsnAlaIleLeuGluArgAlaGlyAsnIleGluLeuTyrThrGlnArg 180
Qy      541 CGGCAAAAGCGGTTTCGCAATCGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTCG 600
Db      181 TyrGlnSerSerPheArgThrLeuGluAsnValLeuAsnPheSerGlnSerGluThrCys 200
Qy      601 CTTAAACGTGAGAACAGGACGAAGCTGTTCAATACGCGAGCATTAACCATCGGAATC 660
Db      201 LysThrThrGluLysSerThrLys---CysThrLeuProGluAlaLeuProSerGluLeu 219
Qy      661 AAGGTGAGCCGCAATGCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Db      220 LysValThrProAspAsnValSerLeuProGlyAlaTrpSerLeuSerSerThrLeuThr 239
Qy      721 GAGATATTCTCTGCAACAGACAGGAAATGCCGAGCGCGGGTGGGGAAGGATCACC 780
Db      240 GluIlePheLeuLeuGlnGluAlaGlnGlyMetProGlnValAlaTrpGlyArgIleThr 259
Qy      781 GATTCACACAGTGAACACCTTGCTAAGTTGTCATACGCGCAATTTTATTGCTACAA 840
Db      260 GlyGluLysGluTrpArgAspLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 279
Qy      841 CGCAGCCGAGAGTGTGCCGCGCGCCACCCCGTTATTGATTTGATCATGCGCAGCG 900
Db      280 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspMetIleAspThrAla 299
Qy      901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGCTGTGACATTACCCACTTCAGTACTG 960
Db      300 LeuLeuThrAsnGlyThrThrGluAsnArgTyrGlyIleLysLeuProValSerLeuLeu 319
Qy      961 TTTATTCCGCGACAGTACTATCTGCAATCTCGCAATCTCGCGCGCAGCTGGAGCTCAACTGG 1020
Db      320 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuSerGlyAlaLeuAspLeuAsnTrp 339
Qy      1021 ACGCTTCCCGCTCAGCCGGATAACACGCGCGAGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
Db      340 SerLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluLysTrp 359
Qy      1081 CGTCGGCTAGCGGATAACAGCCAGTGGATTTCAGGTTTCGGTTCGTCCTCCAGACTTTACAG 1140
Db      360 LysArgThrSerAspAsnThrAspTrpValGlnValSerPheValTyrGlnThrLeuArg 379
```

```
Qy      1141 CAGATCGTGCATAAAACGCCGCTGTCTATTAAATACGCCGCCGAGAGGTGAAACTGACC 1200
Db      380 AspMetArgAspIleGlnProLeuSerLeuGluLysProAlaGlyLysValAspLeuLys 399
Qy      1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTCGAGGTTTACGCAA 1260
Db      400 LeuIleAlaCysGluGluLysAsnSerGlnGlyMetCysSerLeuLysSerPheSerArg 419
Qy      1261 ATCGTGAATGAAGCACGACATACCGCGGTGCAGTTTG 1296
Db      420 LeuIleLysGluIleArgValProGluCysAlaVal 431
```

Search completed: March 11, 2005, 23:36:36
Job time : 277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:55:05 ; Search time 168 Seconds
(without alignments)
1012.944 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	440	4 AAB37892	Aab37892 Escherich
2	2302	100.0	440	5 AAE22836	Aae22836 Escherich
3	2302	100.0	440	5 AAE15806	Aae15806 Escherich
4	2302	100.0	440	6 ADA19446	Ada19446 E. coli B
5	2302	100.0	440	8 ADO50292	Ado50292 Escherich
6	2282	99.1	440	8 ADO50304	Ado50304 Escherich
7	2279	99.0	440	8 ADO50302	Ado50302 Kangaroo
8	2235	97.1	432	3 AAB36257	Aab36257 Lama2/APP
9	2235	97.1	432	3 AAB36261	Aab36261 R15/APP
10	2235	97.1	432	3 AAB36262	Aab36262 SV40/APP
11	2235	97.1	432	3 AAB36259	Aab36259 R15/APP
12	2235	97.1	432	3 AAB36258	Aab36258 R15/APP
13	2235	97.1	432	3 AAB36263	Aab36263 Lama2/APP
14	2235	97.1	432	3 AAB36260	Aab36260 R15/APP
15	2235	97.1	432	3 AAU77775	Aau77775 Phytase a
16	2235	97.1	432	5 AAE15807	Aae15807 Escherich
17	2235	97.1	432	5 ABP53625	Abp53625 Acid phos
18	2235	97.1	432	6 ADA19450	Ada19450 E. coli K
19	2235	97.1	432	6 ABU15427	Abu15427 Protein e
20	2235	97.1	432	7 ADC87745	Adc87745 Escherich
21	2235	97.1	432	8 ADL16129	Adl16129 Escherich
22	2235	97.1	432	8 ADO50298	Ado50298 Escherich
23	2230	96.9	432	4 AAE02631	Aae02631 E. coli a
24	2230	96.9	432	8 ADL16131	Adl16131 Shigella
25	2224	96.6	432	4 AAE02634	Aae02634 E. coli a

ALIGNMENTS

RESULT 1

AAB37892
ID AAB37892 standard; protein; 440 AA.

AC AAB37892;

XX 07-MAR-2001 (first entry)

DE Escherichia coli B phytase enzyme.

XX Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

XX Escherichia coli.

XX WO2000071728-A1.

XX 30-NOV-2000.

XX 25-MAY-2000; 2000WO-US014846.

XX 25-MAY-1999; 99US-00318528.

XX (DIVE-) DIVERSA CORP.

XX Short JM, Kretz KA;

XX WPI; 2001-112081/12.

XX N-PSDB; AAC88885.

Improving the nutritional value of phytate-containing foodstuffs, using phytase enzymes which catalyze the liberation of inorganic phosphate from the phytates.

Claim 1; Fig 1; 147pp; English.

The present sequence is a phytase enzyme from Escherichia coli B. The enzyme catalyses the liberation of inorganic phosphate from the phytate in phytate-containing foodstuffs and can thus be used to improve the nutritional value of phytate rich ingredients

Sequence 440 AA;

Query Match . 100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILPFLSLIPLTPQSAFQAQSELPKLESYVIVSRHGVRAPTKATQLMQDVTDPAMP 60

[illegible]

RESULT 2	
AAE22836	
ID	AAE22836 standard; protein; 440 AA.
XX	
AC	AAE22836;
XX	
DT	21-AUG-2002 (first entry)
XX	
DE	Escherichia coli phytase protein.
XX	
KW	Dietary aid; biocompatible composition; therapeutic; digestive tract;
KW	foodstuff; digestion; phytase; enzyme.
XX	
OS	Escherichia coli.
XX	
PN	WO2001189317-A2.
XX	
PD	29-NOV-2001.
XX	
PF	15-MAY-2001; 2001WO-US015764.
XX	
PR	25-MAY-2000; 2000US-00580937.
XX	
PA	(DIVE-) DIVERSA CORP.
XX	
PI	Short JM, Kretz KA, O'donoghue E;
XX	
DR	WPI; 2002-164149/21.
DR	N-PSDB; AAD36473.
XX	
PT	New dietary aids comprising sustained release biocompatible compositions,
PT	comprise agent that assists in digestion, useful for delivering enzymes,
PT	therapeutics, medicine or agents to an organism.
XX	
PS	Claim 5; Fig 1; 89pp; English.
XX	
CC	The present invention relates to novel dietary aids comprising sustained
CC	release biocompatible composition which comprises an agent (enzymes such
CC	as phytase, amylase, esterase, protease) that assists in digestion. The
CC	biocompatible composition is effective upon oral consumption and release
CC	in the digestive tract of a subject. The dietary aids are useful for

CC	delivering enzymes, therapeutics, medicine and agents to an organism. The
CC	use of enzymes and other agents in digestive aids of livestock or
CC	domesticated animals not only improves the animal's health and life
CC	expectancy but also assists in increasing the health of livestock or in
CC	the production of foodstuffs from livestock. The present sequence is
CC	Escherichia coli phytase protein
XX	
SQ	Sequence 440 AA;
	Query Match 100.0%; Score 2302; DB 5; Length 440;
	Best Local Similarity 100.0%; Pred. No. 2.2e-223;
	Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB	1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY	61 TWPVKLGWLTPRGGEIAYIGHYQORLVADGILLAKKGCPSQGVAIADVDERTKTGE 120
DB	61 TWPVKLGWLTPRGGEIAYIGHYQORLVADGILLAKKGCPSQGVAIADVDERTKTGE 120
QY	121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTCVOLDNANVTDAILSRAGGSIADFTGH 180
DB	121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTCVOLDNANVTDAILSRAGGSIADFTGH 180
QY	181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB	181 RQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY	241 EIFLLQAOQMPFPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIMAA 300
DB	241 EIFLLQAOQMPFPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIMAA 300
QY	301 LTPHPPQKQAYGVTLPSTVSLFIAGHDTNLANLGALELAWTLPGQPDNTPPGGELVFERW 360
DB	301 LTPHPPQKQAYGVTLPSTVSLFIAGHDTNLANLGALELAWTLPGQPDNTPPGGELVFERW 360
QY	361 RRLSDNSQWITQVSLVFTQLQWMDKTPLSLNTPPGVEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB	361 RRLSDNSQWITQVSLVFTQLQWMDKTPLSLNTPPGVEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY	421 IVNEARIPACSLRSHHHHH 440
DB	421 IVNEARIPACSLRSHHHHH 440
RESULT 3	
AAE15806	ID AAE15806 standard; protein; 440 AA.
XX	AAE15806;
XX	26-MAR-2002 (first entry)
DT	Escherichia coli B phytase protein.
DE	
XX	Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW	Gastrointestinal; nutritional value; feed treatment process; therapy;
KW	thermal tolerance; growth performance; alcoholic drink; biopulping;
KW	non-alcoholic drink; biobleaching; B phytase.
XX	
OS	Escherichia coli.
XX	
FN	WO200190333-A2.
XX	
PD	29-NOV-2001.
XX	
PF	24-MAY-2001; 2001WO-US017118.
XX	
PR	25-MAY-2000; 2000US-00580515.
XX	
PA	(DIVE-) DIVERSA CORP.
XX	

PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 XX WPI; 2002-083108/11.
 DR N-PSDB; AAD25460.
 XX
 XX
 XX
 PT New bacterial phytase for e.g. improving the nutritional value of phytate
 PT -containing foodstuffs and subsequently improving the growth performance
 PT of an organism that consumes it, or in treating animal digestive systems.
 XX
 XX
 PS Claim 1; Fig 1; 170pp; English.
 XX
 XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
 CC K12 appA phytase. The enzyme has phytase activity and improved thermal
 CC tolerance when compared with wild-type phytase. It has improved protease
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytate-containing foodstuffs and subsequently
 CC improving the growth performance of an organism that consumes it, in
 CC treating animal digestive systems, in feed treatment processes and for in
 CC vitro purposes related to research, discovery and development. They are
 CC also used for generating recombinant digestive system life forms, for
 CC producing or manufacturing alcoholic and non-alcoholic drinks based on
 CC the use of moulds, grains and/or plants, in biopulping and bio-bleaching
 CC where a reduction in the use of environmentally harmful chemicals that
 CC are traditionally used in the pulp and paper industry is desired and in
 CC the reduction or possible elimination of the need for mineral
 CC supplements, enzymes or therapeutic drugs for animals from the daily feed
 CC thus increasing the amount calories and nutrients present in the feed.
 CC The present sequence is *E. coli* B phytase protein
 XX
 SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 5; Length 440;
 Best Local Similarity 100.0%; Pred. No. 2.2e-223;
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAIIPLFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
 DB 1 MKAIIPLFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
 QY 61 TWPVKGWLTPRGELIAYLGHYQRLVADGLLAKKGCPSQGOVATIAVDERTKTGE 120
 DB 61 TWPVKGWLTPRGELIAYLGHYQRLVADGLLAKKGCPSQGOVATIAVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDANVTDAILSRAGGSIAFTGH 180
 QY 181 ROTAFRELERVLPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELERVLPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLOAQCPMPFGWGRIITDSHOWNTLLSHNAQFVLLQRTPEVARSRAATPLLDLIMAA 300
 DB 241 EIFLLOAQCPMPFGWGRIITDSHOWNTLLSHNAQFVLLQRTPEVARSRAATPLLDLIMAA 300
 QY 301 LTPHPPOKQYGVTLPTSVLFIAGHDNTNLNGLGALSNLTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPPOKQYGVTLPTSVLFIAGHDNTNLNGLGALSNLTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWTVSLVFTQLOMRDKTPLSLNTPPGVEKVLFIAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWTVSLVFTQLOMRDKTPLSLNTPPGVEKVLFIAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSLRSHHHHH 440
 DB 421 IVNEARIPACSLRSHHHHH 440
 RESULT 4
 ADA19446
 ID ADA19446 standard; protein; 440 AA.
 XX
 AC ADA19446;

XX 20-NOV-2003 (first entry)
 XX E. coli B phytase.
 DE
 XX
 KW Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
 KW digestion enhancement; transgenic; thermal tolerance; protease stability.
 XX
 OS *Escherichia coli* B.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT
 XX US2002136754-A1.
 XX 26-SEP-2002.
 XX 24-MAY-2001; 2001US-00866379.
 XX 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 XX (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (DONO/) O' DONOGHUE E.
 PA (MATH/) MATHUR E J.
 XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
 PI O' Donoghue E, Mathur EJ;
 XX WPI; 2003-040002/03.
 XX Isolated *Escherichia coli* polynucleotide encoding a modified phytase
 PT enzyme, useful in the production of animal feed, for improving the
 PT nutritional value of phytate-containing foodstuff and for enhancing
 PT digestion in humans and animals.
 XX
 PS Claim 1; Fig 1; 62pp; English.
 XX The invention relates to an isolated *Escherichia coli* polynucleotide
 CC encoding a phytase enzyme appearing as ADA19450 and having amino acids
 CC modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
 CC Also included the *E. coli* appA gene ADA19449 (or an oligonucleotide
 CC derived from it) or its mutant sequence ADA19452, expression vectors,
 CC host cells, a method of improving nutritional value of a phytate-
 CC containing foodstuff by contacting the phytate-containing foodstuff with
 CC a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
 CC the liberation of inorganic phosphate from the phytate in the phytate-
 CC containing foodstuff), a method to produce an animal feed containing a
 CC microbial phytase (comprising culturing the plant cell, plant part or
 CC plant under conditions where the nucleotide sequence is expressed and
 CC converting the plant cells, plants or plant into a composition for animal
 CC feed), a feed composition for animals (comprising the plant seeds, plant
 CC cells, plant parts or plants in admixture with a phytate-containing
 CC foodstuff), a method to treat a human or an animal able to benefit from
 CC digestive enhancement by the activity of an exogenous phytase enzymes
 CC comprising administering to the human or animal the plant seed, plant
 CC cells, plant parts or plants of a transgenic plant which is modified to
 CC contain an expression system which expresses a nucleotide sequence
 CC encoding a phytase enzyme, a transgenic non-human organism whose genome
 CC comprising a heterologous nucleic acid sequence encoding a polypeptide
 CC having phytase activity. The phytase enzyme is useful for improving the
 CC nutritional value of phytate-containing foodstuff, in the production of
 CC animal feed and for enhancing digestion in humans and animals. The
 CC invented method improves thermal tolerance and protease stability. It
 CC also improves the feeding value of phytate rich ingredients. The present

```
CC sequence represents E. coli B wild-type phytase.
XX
SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPFGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPFGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALNELNWLTPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALNELNWLTPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
AD050292
ID AD050292 standard; protein; 440 AA.
AC AD050292;
XX
XX 29-JUL-2004 (first entry)
XX
XX Escherichia coli B phytase enzyme.
XX
XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytate; animal feed; fish feed; dough; baking; enzyme.
XX
XX Escherichia coli; B.
XX
XX Key Location/Qualifiers
XX Misc-difference 72 /note= "Encoded by CGN"
XX
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX

(SHOR/) SHORT J M.
(KRET/) KRETZ K.
(PA) (GRAY/) GRAY K A.
(PA) (BART/) BARTON N R.
(PA) (GARR/) GARRETT J B.
(PA) (ODON/) O'DONOGHUE E.
(PA) (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
XX
XX MPI; 2004-374952/35.
DR N-PSDB; AD050291.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX
XX Claim 3; SEQ ID NO 2; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Escherichia coli B phytase enzyme.
XX
XX Sequence 440 AA;
Query Match 100.0%; Score 2302; DB 8; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPFGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPFGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALNELNWLTPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALNELNWLTPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440
```

Db 421 IVNEARIPACSLRSHHHH 440

RESULT 6

AD050304

ID AD050304 standard; protein; 440 AA.

XX AC AD050304;

XX DT 29-JUL-2004 (first entry)

XX DE Escherichia coli W phytase 875PH2 mutant enzyme.

XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;

XX KW phytate; animal feed; fish feed; dough; baking; mutant; muten.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT Misc-difference 72 /note= "Encoded by CGN"

FT Misc-difference 160 /note= "E. coli B phytase Ala replaced with Ser"

FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"

FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"

FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"

XX US2004091968-A1.

XX PD 13-MAY-2004.

XX PF 20-JUN-2003; 2003US-00601319.

XX PR 13-AUG-1997; 97US-00910798.

XX PR 01-MAR-1999; 99US-00259214.

XX PR 13-APR-1999; 99US-00291931.

XX PR 25-MAY-1999; 99US-00318528.

XX PR 25-MAY-2000; 2000US-00580515.

XX PR 24-MAY-2001; 2001US-00866379.

XX PA (SHORT) SHORT J M.

XX PA (KRETZ) KRETZ K.

XX PA (GRAY) GRAY K A.

XX PA (BART) BARTON N R.

XX PA (GARR) GARRATT J B.

XX PA (ODON) O'DONOGHUE E.

XX PA (MATH) MATH E J.

XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;

XX Mather EJ;

XX WI; 2004-374952/35.

DR N-PSDB; AD050291, AD050303.

XX PT Producing phytase, involves providing nucleic acid derived from bacteria

XX PT encoding polypeptide having phytase activity, and expressing nucleic acid

XX PT in yeast.

XX PS Disclosure; Page; 74pp; English.

XX CC The invention relates to a method for producing phytase that involves

XX CC providing a nucleic acid encoding phytase derived from a bacteria, and

XX CC expressing the nucleic acid in a yeast under conditions that allow

XX CC expression of the enzyme in the yeast. The invention also relates to

XX CC modified phytase enzyme which has improved thermal tolerance and protease

XX CC stability at low pH. The phytase enzyme can be used in foodstuffs to

XX CC improve the feeding value of phytate rich ingredients, and in diet of

XX CC numerous animals including mammals, fowls and fishes, commercially

XX CC significant mammals such as pigs, goats, laboratory rodents, commercially

XX CC significant avian species such as chicken, ducks, doves, parrot, etc.,

CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,

CC etc., in dough making and baking, in dietary aids for animals. The method

CC provides easy manufacture of the active ingredient loaded biocompatible

CC composition, higher yields and loading efficiency. The phytase

CC incorporated in to the dietary aid is safe for animals. The present

CC sequence is Escherichia coli W phytase mutant enzyme. Note: The present

CC sequence is not shown in the specification but has been derived from

CC Escherichia coli B phytase AD050292.

XX SQ Sequence 440 AA;

Query Match 99.1%; Score 2282; DB 8; Length 440;

Best Local Similarity 99.1%; Pred. No. 2.4e-221;

Matches 436; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

DB 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPCQSQVIAIADVDERTRKTGE 120

DB 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPCQSQVIAIADVDERTRKTGE 120

QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

QY 181 ROTAFRELERVLNFPQSNICLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

DB 181 ROTAFRELERVLNFPQSNICLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOQAQMPGPGWGRITDSHWNLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

DB 241 EIFLLOQAQMPGPGWGRITDSHWNLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPKQAYGVTLPTSVLFTAGHDNLANLGGALELANWTLPGQDNTPPGGELVPERW 360

DB 301 LTPHPKQAYGVTLPTSVLFTAGHDNLANLGGALELANWTLPGQDNTPPGGELVPERW 360

QY 361 RRLSDNSQMIQSVLFTQQLQMRDTPSLINTPPGVEVKLTLAGCEERNAGCMCSLAGFTQ 420

DB 361 RRLSDNSQMIQSVLFTQQLQMRDTPSLINTPPGVEVKLTLAGCEERNAGCMCSLAGFTQ 420

QY 421 IVNEARIPACSLRSHHHH 440

DB 421 IVNEARIPACSLRSHHHH 440

RESULT 7

AD050302

ID AD050302 standard; protein; 440 AA.

XX AC AD050302;

XX DT 29-JUL-2004 (first entry)

XX DE Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.

XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;

XX KW phytate; animal feed; fish feed; dough; baking; mutant; muten.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT Misc-difference 72 /note= "Encoded by CGN"

FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"

FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"

FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"

FT Misc-difference 312

/note= "E. coli B phytase Gly replaced with Ser"	
US2004091968-A1.	
13-MAY-2004.	
20-JUN-2003; 2003US-00601319.	
13-AUG-1997; 97US-00910798.	
01-MAR-1999; 99US-00259214.	
13-APR-1999; 99US-00291931.	
25-MAY-1999; 99US-00318528.	
25-MAY-2000; 2000US-00580515.	
24-MAY-2001; 2001US-00866379.	
(SHOR/) SHORT J M.	
(KRET/) KRETZ K.	
(GRAY/) GRAY K A.	
(BART/) BARTON N R.	
(GARR/) GARRETT J B.	
(ODON/) O'DONOGHUE E.	
(MATH/) MATHER E J.	
Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E; Mather EJ;	
WPI; 2004-374952/35.	
N-PSDB; ADOS0301.	
Producing phytase, involves providing nucleic acid derived from bacteria encoding polypeptide having phytase activity, and expressing nucleic acid in yeast.	
Disclosure; Page; 74pp; English.	
The invention relates to a method for producing phytase that involves providing a nucleic acid encoding phytase derived from a bacteria, and expressing the nucleic acid in a yeast under conditions that allow expression of the enzyme in the yeast. The invention also relates to modified phytase enzyme which has improved thermal tolerance and protease stability at low pH. The phytase enzyme can be used in feedstuffs to improve the feeding value of phytate rich ingredients, and in diet of numerous animals including mammals, fowls and fishes, commercially significant mammals such as pigs, goats, laboratory rodents, commercially significant avian species such as chicken, ducks, doves, parrot, etc., commercially farmed fish such as guppy, zebrafish, molly, swordtail, etc., in dough making and baking, in dietary aids for animals. The method provides easy manufacture of the active ingredient loaded biocompatible composition, higher yields and loading efficiency. The phytase incorporated in to the dietary aid is safe for animals. The present sequence is Kangaroo rat Escherichia coli phytase mutant enzyme. Note: The present sequence is not shown in the specification but has been derived from Escherichia coli B phytase sequence ADOS0292.	
Sequence 440 AA;	
Query Match	99.08; Score 2279; DB 8; Length 440;
Best Local Similarity	99.18; Pred. No. 4.7e-221;
Matches 436; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
Qy	1 MKAILIFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db	1 MKAILIFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Qy	61 TWPVKLWLTFRGELLAYLGHYORQLVADGLLAKGCPQSGQVAILADVDERTRKTGE 120
Db	61 TWPVKLWLTFRGELLAYLGHYORQLVADGLLAKGCPQSGQVAILADVDERTRKTGE 120
Qy	121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDNANVTDAILSRAGGSFADPTGH 180
Db	121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDNANVTDAILSRAGGSFADPTGH 180
Qy	181 RQTAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Db	181 RQTAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy	241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRAATPLDLIMAA 300
Db	241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRAATPLDLIKTA 300
Qy	301 LTPHPQKQAYGVTLPTSVLFIAGHDNTNLNGLGALNLTWTLFGQPDNTPPGGELVFERW 360
Db	301 LTPHPQKQAYSVTLPTSVLFIAGHDNTNLNGLGALNLTWTLFGQPDNTPPGGELVFERW 360
Qy	361 RRLSDNSQWITQVSLVFTQLQOMRDKTPTLSINTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db	361 RRLSDNSQWITQVSLVFTQLQOMRDKTPTLSINTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy	421 IVNEARIPACSLRSHHHHH 440
Db	421 IVNEARIPACSLRSHHHHH 440
RESULT 8	
AAB36257	
ID	AAB36257 standard; protein; 432 AA.
XX	AC
XX	AAB36257;
XX	12-SEP-2003 (revised)
DT	20-FEB-2001 (first entry)
XX	Lama2/APPA plasmid translated sequence.
DE	Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW	environmental pollution; pig.
XX	Mus musculus.
OS	Escherichia coli.
OS	Chimeric.
PN	WO200064247-A1.
XX	02-NOV-2000.
PD	20-APR-2000; 2000WO-CA000430.
PF	23-APR-1999; 99US-0130508P.
XX	(UYGU-) UNIV GUELPH.
XX	Forsberg CW, Golovan S, Phillips JP;
XX	WPI; 2000-687245/67.
DR	N-PSDB; AAC68294.
XX	Transgenic non-human animal for gastrointestinal tract specific
PT	expression of a protein, preferably phytase, comprises a nucleic acid
PT	sequence including a heterologous transgene construct encoding the
PT	protein.
XX	Disclosure; Fig 5; 152pp; English.
CC	The present invention provides transgenic animals which produce desired
CC	proteins, in this case pigs which expresses phytase in the salivary
CC	gland. Low phytase production levels result in phytate in the diet being
CC	excreted and causing phosphorus contamination in water, as well as
CC	reducing the growth of animals. The invention provides a number of
CC	transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC	on 12-SEP-2003 to standardise OS field)
XX	Sequence 432 AA;
SQ	Query Match
	Best Local Similarity
	Matches 430; Conservative
	97.18; Score 2235; DB 3; Length 432;
	99.58; Pred. No. 1.3e-216;
	0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
 DB 1 MKAIIIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAVIADVDERTRKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAVIADVDERTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
 QY 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
 DB 241 EIFLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
 QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 9

AAB36261
 ID AAB36261 standard; protein; 432 AA.

XX AC AAB36261;

DT 12-SEP-2003 (revised)

DT 20-FEB-2001 (first entry)

DE R15/APPA plasmid translated sequence.

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.

OS Rattus sp.

OS Escherichia coli.

OS Chimeric.

XX PN WO200064247-A1.

XX PD 02-NOV-2000.

XX PF 20-APR-2000; 2000WO-CA000430.

XX PR 23-APR-1999; 99US-0130508P.

XX PA (UYGU-) UNIV GUELPH.

XX PI Forsberg CW, Golovan S, Phillips JP;

XX DR WPI; 2000-687245/67.

XX DR N-PSDB; AAC68298.

XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.

XX PS Disclosure; Fig 21; 152pp; English.

XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 432 AA;

Query Match 97.1%; Score 2235; DB 3; Length 432;

Best Local Similarity 99.5%; Pred. No. 1.3e-216;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60

DB 1 MKAIIIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60

QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAVIADVDERTRKTGE 120

DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAVIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180

DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180

QY 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

DB 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300

DB 241 EIFLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIKTA 300

QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420

DB 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420

QY 421 IVNEARIPACSL 432

DB 421 IVNEARIPACSL 432

RESULT 10

AAB36262

ID AAB36262 standard; protein; 432 AA.

XX AC AAB36262;

DT 12-SEP-2003 (revised)

DT 20-FEB-2001 (first entry)

DE SV40/APPA plasmid translated sequence.

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.

XX OS Simian virus 40.

OS Escherichia coli.

OS Chimeric.

XX PN WO200064247-A1.

XX PD 02-NOV-2000.

XX PF 20-APR-2000; 2000WO-CA000430.

XX PR 23-APR-1999; 99US-0130508P.

XX (UYGU-) UNIV GUELPH.
 XX Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 XX N-PSDB; AAC68299.
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX Disclosure; Fig 22; 152pp; English.
 XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX Sequence 432 AA;
 SQ Query Match 97.1%; Score 2235; DB 3; Length 432;
 Best Local Similarity 99.5%; Pred. No. 1.3e-216; Indels 0; Gaps 0;
 Matches 430; Conservative 0; Mismatches 2;
 QY 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 DB 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGILLAKKGCPSQGOVAILIADVDRTRKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGILLAKKGCPSQGOVAILIADVDRTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIADFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIADFTGH 180
 QY 181 QTAFARELRLVNFPPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 QTAFARELRLVNFPPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
 DB 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
 QY 301 LTPHPKQKQAYGVTLPSTVSLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPKQKQAYGVTLPSTVSLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432
 RESULT 11
 AAB36259 standard; protein; 432 AA.
 ID AAB36259
 XX AC AAB36259;
 XX AC AAB36259;
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX R15/APPA plasmid translated sequence.
 DE XX
 XX

KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX WO200064247-A1.
 XX 02-NOV-2000.
 XX 20-APR-2000; 2000WO-CA000430.
 XX 23-APR-1999; 99US-0130508P.
 XX (UYGU-) UNIV GUELPH.
 XX Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 XX N-PSDB; AAC68299.
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX Disclosure; Fig 19; 152pp; English.
 XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX Sequence 432 AA;
 SQ Query Match 97.1%; Score 2235; DB 3; Length 432;
 Best Local Similarity 99.5%; Pred. No. 1.3e-216; Indels 0; Gaps 0;
 Matches 430; Conservative 0; Mismatches 2;
 QY 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 DB 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGILLAKKGCPSQGOVAILIADVDRTRKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGILLAKKGCPSQGOVAILIADVDRTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIADFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIADFTGH 180
 QY 181 QTAFARELRLVNFPPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 QTAFARELRLVNFPPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
 DB 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
 QY 301 LTPHPKQKQAYGVTLPSTVSLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPKQKQAYGVTLPSTVSLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

```

Db      421  IVNEARIPACSL 432

RESULT 12
AAB36258
ID  AAB36258 standard; protein; 432 AA.
AC  AAB36258;
XX
XX  12-SEP-2003 (revised)
DT  20-FEB-2001 (first entry)
XX
DE  R15/APPA plasmid translated sequence.
XX
XX  Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW  environmental pollution; pig.
XX
OS  Rattus sp.
OS  Escherichia coli.
XX
XX  WO200064247-A1.
XX
XX  02-NOV-2000.
PD
XX
XX  20-APR-2000; 2000WO-CA000430.
PF
XX
XX  23-APR-1999; 99US-0130508P.
PR
XX
XX  (UYGU-) UNIV GUELPH.
PA
XX
XX  Forsberg CW, Golovan S, Phillips JP;
PI
XX  WPI; 2000-687245/67.
DR
XX  N-PSDB; AAC68295.
XX
XX  Transgenic non-human animal for gastrointestinal tract specific
PT  expression of a protein, preferably phytase, comprises a nucleic acid
PT  sequence including a heterologous transgene construct encoding the
PT  protein.
XX
XX  Disclosure; Fig 18; 152pp; English.
PS
XX
XX  The present invention provides transgenic animals which produce desired
CC  proteins, in this case pigs which expresses phytase in the salivary
CC  gland. Low phytase production levels result in phytate in the diet being
CC  excreted and causing phosphorus contamination in water, as well as
CC  reducing the growth of animals. The invention provides a number of
CC  transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC  on 12-SEP-2003 to standardise OS field)
XX
SQ  Sequence 432 AA;

Query Match      97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db      1  MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Qy      61  TWPVKLGLWTPRGGLIAYLGHYQORLVADGLLAKKGCQSGQVAILADVDERTRKTGE 120
Db      61  TWPVKLGLWTPRGGLIAYLGHYQORLVADGLLAKKGCQSGQVAILADVDERTRKTGE 120
Qy      121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db      121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy      181  QTAFARELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181  QTAFARELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy      241  EIFLQQAQGMPEPGWGGRITDTSQWNTLLSLHNAQFYLLQRTBEVARSRATPLLDLIMAA 300
Db      241  EIFLQQAQGMPEPGWGGRITDTSQWNTLLSLHNAQFYLLQRTBEVARSRATPLLDLIMAA 300
Qy      301  LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGSELVERW 360
Db      301  LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGSELVERW 360
Qy      361  RRLSDNSQWITQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db      361  RRLSDNSQWITQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy      421  IVNEARIPACSL 432
Db      421  IVNEARIPACSL 432

RESULT 13
AAB36263
ID  AAB36263 standard; protein; 432 AA.
XX
XX  AAB36263;
XX
XX  12-SEP-2003 (revised)
DT  20-FEB-2001 (first entry)
XX
XX  Lama2/APPA plasmid translated sequence.
DE
XX
XX  Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW  environmental pollution; pig.
XX
XX  Mus musculus.
OS  Escherichia coli.
OS  Chimeric.
XX
XX  WO200064247-A1.
PN
XX
XX  02-NOV-2000.
PD
XX
XX  20-APR-2000; 2000WO-CA000430.
PF
XX
XX  23-APR-1999; 99US-0130508P.
PR
XX
XX  (UYGU-) UNIV GUELPH.
PA
XX
XX  Forsberg CW, Golovan S, Phillips JP;
PI
XX
XX  WPI; 2000-687245/67.
DR
XX  N-PSDB; AAC68300.
XX
XX  Transgenic non-human animal for gastrointestinal tract specific
PT  expression of a protein, preferably phytase, comprises a nucleic acid
PT  sequence including a heterologous transgene construct encoding the
PT  protein.
XX
XX  Disclosure; Fig 23; 152pp; English.
PS
XX
XX  The present invention provides transgenic animals which produce desired
CC  proteins, in this case pigs which expresses phytase in the salivary
CC  gland. Low phytase production levels result in phytate in the diet being
CC  excreted and causing phosphorus contamination in water, as well as
CC  reducing the growth of animals. The invention provides a number of
CC  transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC  on 12-SEP-2003 to standardise OS field)
XX
SQ  Sequence 432 AA;

Query Match      97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db      1  MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Qy      61  TWPVKLGLWTPRGGLIAYLGHYQORLVADGLLAKKGCQSGQVAILADVDERTRKTGE 120
Db      61  TWPVKLGLWTPRGGLIAYLGHYQORLVADGLLAKKGCQSGQVAILADVDERTRKTGE 120
Qy      121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db      121  AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy      181  QTAFARELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181  QTAFARELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240

```

Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120
 Db 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILLSRAGGSIADFTGH 180
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILLSRAGGSIADFTGH 180
 QY 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 Db 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 QY 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWIOVSLVFTQLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 Db 361 RRLSDNSQWIOVSLVFTQLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 Db 421 IVNEARIPACSL 432

RESULT 14

AAB36260 standard; protein; 432 AA.
 ID AAB36260 standard; protein; 432 AA.
 XX

AC AAB36260;
 XX

DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)

XX R15/APPA plasmid translated sequence.
 DE
 XX

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX

XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX

XX WO200064247-A1.
 XX

XX 02-NOV-2000.
 XX

XX 20-APR-2000; 2000WO-CA000430.
 XX

XX 23-APR-1999; 99US-0130508P.
 XX

XX (UYGU-) UNIV GUELPH.
 XX

XX Forsberg CW, Golovan S, Phillips JP;
 PI

XX WPI; 2000-687245/67.
 DR

XX N-PSDB; AAC68297.
 DR

XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX

XX Disclosure; Fig 20; 152pp; English.
 PS

XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC

CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 432 AA;

Query Match 97.1%; Score 2235; DB 3; Length 432;
 Best Local Similarity 99.5%; Pred. No. 1.3e-216;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60

QY 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120
 Db 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILLSRAGGSIADFTGH 180
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILLSRAGGSIADFTGH 180

QY 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 Db 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWIOVSLVFTQLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 Db 361 RRLSDNSQWIOVSLVFTQLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
 Db 421 IVNEARIPACSL 432

RESULT 15

AAU77775
 ID AAU77775 standard; protein; 432 AA.
 XX

AC AAU77775;
 XX

DT 05-JUN-2002 (first entry)
 XX

XX Phytase associated protein.
 DE

XX Phytase.
 KW

XX Unidentified.
 OS

XX KR99086028-A.
 PN

XX 15-DEC-1999.
 PD

XX 25-MAY-1998; 98KR-00018810.
 PF

XX 25-MAY-1998; 98KR-00018810.
 PR

XX (WOJ-) WOJIN CO LTD.
 PA

XX Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 PI

XX WPI; 2000-645078/62.
 DR

XX N-PSDB; ABK12514.
 DR

XX Novel phytase gene, recombinant phytase and usage thereof.
XX Disclosure; Fig 3; 10pp; Korean.
XX The invention relates to a novel phytase gene, a recombinant phytase gene
CC and their uses. This is the amino acid sequence of the phytase associated
CC protein described in the invention
XX
SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216; Mismatches 0; Indels 0; Gaps 0;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGWLTPRGGLIAYLGHYQRLVADGLLAKKGCPSGQVAILIADVDERTRKTGE 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 TWPVKLGWLTPRGGLIAYLGHYQRLVADGLLAKKGCPSGQVAILIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSSIADFTGH 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSSIADFTGH 180
QY 181 ROTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 ROTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 EIFLLQQAQGMPEPGWGRIITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
QY 301 LTPHPKQKAYGVTLPSTVLFIAHGDFTNLNLCGALNLNWTLPQPDNTPPGGELVPERW 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 LTPHPKQKAYGVTLPSTVLFIAHGDFTNLNLCGALNLNWTLPQPDNTPPGGELVPERW 360
QY 361 RRLSDNSQWLVFOTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 RRLSDNSQWLVFOTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB ||||||||||||
421 IVNEARIPACSL 432

Search completed: March 11, 2005, 23:04:48
Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:59:36 ; Search time 43 Seconds
(without alignments)
763.851 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACLSRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	440	3	US-09-259-214-2
2	2302	100.0	440	3	US-09-318-528-2
3	2302	100.0	440	3	US-09-291-931-2
4	2302	100.0	440	4	US-09-580-515-2
5	2221	96.5	433	4	US-09-540-149A-1
6	2194.5	95.3	423	2	US-08-910-798-2
7	528	22.9	421	4	US-08-489-039A-7512
8	376	16.3	522	4	US-09-489-039A-13501
9	119	5.2	439	3	US-09-044-718-3
10	119	5.2	439	4	US-10-062-848-3
11	119	5.2	440	4	US-09-684-855-106
12	119	5.2	440	4	US-09-684-855-128
13	119	5.2	440	4	US-09-684-855-151
14	119	5.2	440	4	US-09-488-265B-6
15	119	5.2	449	3	US-09-044-718-12
16	119	5.2	449	4	US-10-062-848-12
17	119	5.2	465	3	US-08-868-435-33
18	119	5.2	465	3	US-08-744-231-33
19	119	5.2	465	3	US-08-044-718-78
20	119	5.2	465	4	US-09-636-499-6
21	119	5.2	465	4	US-09-273-871A-8
22	119	5.2	465	4	US-10-083-452-8
23	119	5.2	465	4	US-09-635-504-33
24	119	5.2	465	4	US-10-062-848-78
25	119	5.2	440	4	US-09-684-855-107
26	118	5.1	440	4	US-09-684-855-129
27	118	5.1	440	4	US-09-684-855-152

28	118	5.1	440	4	US-09-488-265B-7	Sequence 7, Appli
29	117	5.1	386	1	US-08-758-213-1	Sequence 1, Appli
30	117	5.1	386	2	US-08-692-787-48	Sequence 48, Appl
31	117	5.1	386	3	US-09-097-199-48	Sequence 48, Appl
32	117	5.1	386	4	US-09-949-016-6022	Sequence 6022, Ap
33	117	5.1	515	2	US-09-146-283-2	Sequence 2, Appli
34	117	5.1	515	3	US-08-579-823A-2	Sequence 2, Appli
35	117	5.1	515	3	US-09-344-195-2	Sequence 2, Appli
36	114.5	5.0	326	4	US-09-270-767-41623	Sequence 41623, A
37	111	4.8	440	4	US-09-684-855-108	Sequence 108, App
38	111	4.8	440	4	US-09-684-855-130	Sequence 130, App
39	111	4.8	440	4	US-09-684-855-153	Sequence 153, App
40	111	4.8	440	4	US-09-488-265B-8	Sequence 8, Appli
41	111	4.8	465	3	US-09-044-718-80	Sequence 80, Appl
42	111	4.8	465	4	US-10-062-848-80	Sequence 80, Appl
43	111	4.8	467	4	US-09-684-855-165	Sequence 165, App
44	111	4.8	467	4	US-09-488-265B-33	Sequence 33, Appl
45	110	4.8	440	4	US-09-684-855-109	Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Query Match					100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity					100.0%; Pred. No. 1e-233;
Matches					440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	MKAILPFLSLIPLTPQSAFQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP	60		
Db	1	MKAILPFLSLIPLTPQSAFQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP	60		
Qy	61	TWPKVGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE	120		
Db	61	TWPKVGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE	120		
Qy	121	AFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH	180		
Db	121	AFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH	180		
Qy	181	QTAFRELERLVNFPQSNCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT	240		
Db	181	QTAFRELERLVNFPQSNCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT	240		
Qy	241	EIFLLOAQMPGPGWGRITDHSQWNTLLSIHNAQFYLLQRTPEVARSRATPLLDLIMAA	300		
Db	241	EIFLLOAQMPGPGWGRITDHSQWNTLLSIHNAQFYLLQRTPEVARSRATPLLDLIMAA	300		
Qy	301	LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGSALELNWTLPGQPDNTPPGSELVERW	360		
Db	301	LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGSALELNWTLPGQPDNTPPGSELVERW	360		
Qy	361	RLSDNSQWTVQSLVFTQLQOMRDKTPLSLNTPPGVKLTLACGEERNAQMCSLAGTQ	420		

Db 361 RRLSDNSQWIVQSLVFTQLQMDKPTLSLNTPPGEVKLTLAGCERNAGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHH 440
Db 421 IVNEARIPACSLRSHHHH 440

RESULT 2
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

Query Match 100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 1e-233;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLIPTPSAFQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Db 1 MKAILPFLSLLIPTPSAFQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60

Qy 61 TWPVKLWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

Qy 181 QTAFARELRLVNFPPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELRLVNFPPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLLQQAQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNGLGALFNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNGLGALFNWTLPGQPDNTPPGGELVPERW 360

Qy 361 RRLSDNSQWIVQSLVFTQLQMDKPTLSLNTPPGEVKLTLAGCERNAGMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSLVFTQLQMDKPTLSLNTPPGEVKLTLAGCERNAGMCSLAGFTQ 420

Qy 421 IVNEARIPACSLRSHHHH 440
Db 421 IVNEARIPACSLRSHHHH 440

RESULT 3
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:

; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

Query Match 100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 1e-233;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLIPTPSAFQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Db 1 MKAILPFLSLLIPTPSAFQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60

Qy 61 TWPVKLWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

Qy 181 QTAFARELRLVNFPPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELRLVNFPPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLLQQAQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNGLGALFNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNGLGALFNWTLPGQPDNTPPGGELVPERW 360

Qy 361 RRLSDNSQWIVQSLVFTQLQMDKPTLSLNTPPGEVKLTLAGCERNAGMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSLVFTQLQMDKPTLSLNTPPGEVKLTLAGCERNAGMCSLAGFTQ 420

Qy 421 IVNEARIPACSLRSHHHH 440
Db 421 IVNEARIPACSLRSHHHH 440

RESULT 4
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PR1
; ORGANISM: Escherichia coli
US-09-580-515-2

Query Match
Best Local Similarity 100.0%; Score 2302; DB 4; Length 440;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTATQMLQMDVTPDAMP 60
Db 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTATQMLQMDVTPDAMP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQMGPEPGWGRIITDSSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLOAQMGPEPGWGRIITDSSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLWLGALLENMTLPGQDNTTPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLWLGALLENMTLPGQDNTTPGGELVFERW 360
Qy 361 RLSDNSQWLOVSLVFTLOQMRDTPLSLNTTPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RLSDNSQWLOVSLVFTLOQMRDTPLSLNTTPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRHHHHH 440
Db 421 IVNEARIPACSLRHHHHH 440

RESULT 5
US-09-540-149A-1
; Sequence 1, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540.149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 2000-03-31
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PR1
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match
Best Local Similarity 96.5%; Score 2221; DB 4; Length 433;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTATQMLQMDVTPDAMP 60
```

```
Db 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTATQMLQMDVTPDAMP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQMGPEPGWGRIITDSSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLOAQMGPEPGWGRIITDSSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLWLGALLENMTLPGQDNTTPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLWLGALLENMTLPGQDNTTPGGELVFERW 360
Qy 361 RLSDNSQWLOVSLVFTLOQMRDTPLSLNTTPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RLSDNSQWLOVSLVFTLOQMRDTPLSLNTTPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 6
US-08-910-798-2
; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.798
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-798-2

Query Match 95.3%; Score 2194.5; DB 2; Length 423;
Best Local Similarity 96.1%; Pred. No. 2.1e-222;
```


US-09-044-718-3
; Sequence 3, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-3

Query Match 5.2%; Score 119; DB 3; Length 439;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 26 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKSKYKLVTAIQANATDF 82

Qy 59 -----WPTWPKLGM--LTPRGGLIAYLG--HYQRQLVADGLLAKKGCPOSGQVAII 108
Db 83 KGKFAFLKTYNTLGGADLLTPFGQQVNSGKIFQYQYKALARSVP-----FIR 132

Qy 109 ADVDERTRKTGEAPAAGL-----APDCAITVHTQADTSSPDPPLNPLKTV 154
Db 133 ASGSDRVIASGEKFIQFQQAADPCATNRAAPASVII-PESETFN-----NTLDHGV 186

Qy 155 COLDNANVTDAILSRAGGSI-ADPTGHRQTAFARELERVLNFPQSNCLCKEKQDECSLT 213
Db 187 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 225

Qy 214 QALPSEL--KVSADNVSLTCAVSLASMLTEIFLLQQAQGMPEPCWGRIITDSHOWNTLLSL 271
Db 226 DEDVSLMDMCSFDTVARTSDASQLSPFCOLFTHNE-----W-----KKYNYLOSL 271

Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 272 --GKYGYGAGNPLGPAQGIQFTNELIARLTRSPVQDHTSTNSTLVSNPATFPLNATMYV 329

Qy 324 --GHDNTLANLGGALELNTLPGOPDNTPPGELVFERWRRLSD-NSQWI---QVSLVFQ 377
Db 330 DFSDHNSMVSIFPAL-----GLYNGTEPLSRSTSVEKELDGYASASVWVPPGARAYFE 382

Qy 378 TLQMRDKTPLSLNTPPEVKLTLAGCEERNAQGMCSLAGFTQIVNEAR 426
Db 383 TMQCKSEKPL-VRALINDRVPLHGCDV-DKLGRCCKLNDFVKGLSWAR 429

RESULT 10
US-10-062-848-3
; Sequence 3, Application US/10062848
; Patent No. 6734004
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES

FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/10/062,848
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-062-848-3

Query Match 5.2%; Score 119; DB 4; Length 439;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 26 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKSKYKLVTAIQANATDF 82

Qy 59 -----WPTWPKLGM--LTPRGGLIAYLG--HYQRQLVADGLLAKKGCPOSGQVAII 108
Db 83 KGKFAFLKTYNTLGGADLLTPFGQQVNSGKIFQYQYKALARSVP-----FIR 132

Qy 109 ADVDERTRKTGEAPAAGL-----APDCAITVHTQADTSSPDPPLNPLKTV 154
Db 133 ASGSDRVIASGEKFIQFQQAADPCATNRAAPASVII-PESETFN-----NTLDHGV 186

Qy 155 COLDNANVTDAILSRAGGSI-ADPTGHRQTAFARELERVLNFPQSNCLCKEKQDECSLT 213
Db 187 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 225

Qy 214 QALPSEL--KVSADNVSLTCAVSLASMLTEIFLLQQAQGMPEPCWGRIITDSHOWNTLLSL 271
Db 226 DEDVSLMDMCSFDTVARTSDASQLSPFCOLFTHNE-----W-----KKYNYLOSL 271

Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 272 --GKYGYGAGNPLGPAQGIQFTNELIARLTRSPVQDHTSTNSTLVSNPATFPLNATMYV 329

Qy 324 --GHDNTLANLGGALELNTLPGOPDNTPPGELVFERWRRLSD-NSQWI---QVSLVFQ 377
Db 330 DFSDHNSMVSIFPAL-----GLYNGTEPLSRSTSVEKELDGYASASVWVPPGARAYFE 382

Qy 378 TLQMRDKTPLSLNTPPEVKLTLAGCEERNAQGMCSLAGFTQIVNEAR 426
Db 383 TMQCKSEKPL-VRALINDRVPLHGCDV-DKLGRCCKLNDFVKGLSWAR 429

RESULT 11
US-09-684-855-106
; Sequence 106, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 13073
US-09-684-855-106

Query Match	5.2%;	Score 119;	DB 4;	Length 440;	
Best Local Similarity	21.7%;	Pred. No. 0.0017;			
Matches 102;	Conservative 70;	Mismatches 183;	Indels 114;	Gaps 24;	
Qy	7	PFLSLIPLTPQSAFAQSEPELKESVWVSRHGVRAPT	-----KATOLMODVTPDA---	58	
Db	27	PFPSLEDELSVSSKLPK	---DCRITLVQVLSRHGARYPTSSKKYKKLVTAIQANATDF	83	
Qy	59	-----WPTWPVKLGW--LTPRGGLIAYLG--HYQORQLVADGLLAKKGCPCSQQVAII	108		
Db	84	KGKFAFLKTYNTLGGADLLTPFGEQQLVNSGKFKYQRYKALARSVVP	-----FIR 133		
Qy	109	ADVDERTRKTGEFAAGL	-----APCAITVHTQADTSSPDPFLNPLKTGV	154	
Db	134	ASGSDRVIASGEKFIETGQQAKLADPGATNRAAPAI	SVII--PESETFN-----NTLDHGV	187	
Qy	155	CQLDANVTDAILSRAGSI--ADFTGHRQTAFRELERNVFN	PNFQSNCLCKREKQDESCSLT	213	
Db	188	C-----TKFEASQLGDEVAANFT	-----ALFAP--DIRARAEKHLPGVTLT	226	
Qy	214	QALPSEL--KVSADNVSLTGAVSLASMLTEIFLQQAQGMPE	PGWGRITDTSHOWNTLLSL	271	
Db	227	DEDVSLMDMCSFDTVARTSDASQLSPFCQLFTHNE	-----W-----KKNYLQSL	272	
Qy	324	--GHDTNLNLANLGGALELWTLPGQDNTPPGGELVFER	WRRLSD--NSOWI---QVSLVFQ	377	
Db	331	DFSHDMSVSIFFAL	-----GLYNGTEPLSRTSVESAKELDGV	SASWVVPFGARAYFE	383
Qy	378	TLOQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMC	SLAGFTQIYNEAR	426	
Db	384	TMQCKSEKPL--VRALINDRVVPLHGCDV--DKLGRCK	LNDVFKGLSWAR	430	
RESULT 12					
US-09-684-855-128					
Sequence 128,	Application US/09684855				
Patent No. 6599735					
GENERAL INFORMATION:					
APPLICANT:	F. Hoffmann-La Roche AG				
TITLE OF INVENTION:	CONTINUOUS FERMENTATION PROCESS				
FILE REFERENCE:	C38435/111692				
CURRENT APPLICATION NUMBER:	US/09/684,855				
CURRENT FILING DATE:	2000-10-06				
PRIOR APPLICATION NUMBER:	EP 00121663.9				
PRIOR FILING DATE:	2000-10-04				
PRIOR FILING DATE:	2000-10-04				
PRIOR FILING DATE:	1999-10-11				
NUMBER OF SEQ ID NOS:	169				
SOFTWARE:	PatentIn Ver. 2.0				
SEQ ID NO 128					
LENGTH:	440				
TYPE:	PRT				
ORGANISM:	A. fumigatus 13073				
US-09-684-855-128					
Query Match	5.2%;	Score 119;	DB 4;	Length 440;	
Best Local Similarity	21.7%;	Pred. No. 0.0017;			
Matches 102;	Conservative 70;	Mismatches 183;	Indels 114;	Gaps 24;	
Qy	7	PFLSLIPLTPQSAFAQSEPELKESVWVSRHGVRAPT	-----KATOLMODVTPDA---	58	
Db	27	PFPSLEDELSVSSKLPK	---DCRITLVQVLSRHGARYPTSSKKYKKLVTAIQANATDF	83	
Qy	59	-----WPTWPVKLGW--LTPRGGLIAYLG--HYQORQLVADGLLAKKGCPCSQQVAII	108		
Db	84	KGKFAFLKTYNTLGGADLLTPFGEQQLVNSGKFKYQRYKALARSVVP	-----FIR 133		
Qy	109	ADVDERTRKTGEFAAGL	-----APCAITVHTQADTSSPDPFLNPLKTGV	154	
Db	134	ASGSDRVIASGEKFIETGQQAKLADPGATNRAAPAI	SVII--PESETFN-----NTLDHGV	187	
Qy	155	CQLDANVTDAILSRAGSI--ADFTGHRQTAFRELERNVFN	PNFQSNCLCKREKQDESCSLT	213	
Db	188	C-----TKFEASQLGDEVAANFT	-----ALFAP--DIRARAEKHLPGVTLT	226	
Qy	214	QALPSEL--KVSADNVSLTGAVSLASMLTEIFLQQAQGMPE	PGWGRITDTSHOWNTLLSL	271	
Db	227	DEDVSLMDMCSFDTVARTSDASQLSPFCQLFTHNE	-----W-----KKNYLQSL	272	
Qy	324	--GHDTNLNLANLGGALELWTLPGQDNTPPGGELVFER	WRRLSD--NSOWI---QVSLVFQ	377	
Db	331	DFSHDMSVSIFFAL	-----GLYNGTEPLSRTSVESAKELDGV	SASWVVPFGARAYFE	383
Qy	378	TLOQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMC	SLAGFTQIYNEAR	426	
Db	384	TMQCKSEKPL--VRALINDRVVPLHGCDV--DKLGRCK	LNDVFKGLSWAR	430	
RESULT 13					
US-09-684-855-151					
Sequence 151,	Application US/09684855				
Patent No. 6599735					
GENERAL INFORMATION:					
APPLICANT:	F. Hoffmann-La Roche AG				
TITLE OF INVENTION:	CONTINUOUS FERMENTATION PROCESS				
FILE REFERENCE:	C38435/111692				
CURRENT APPLICATION NUMBER:	US/09/684,855				
CURRENT FILING DATE:	2000-10-06				
PRIOR APPLICATION NUMBER:	EP 00121663.9				
PRIOR FILING DATE:	2000-10-04				
PRIOR APPLICATION NUMBER:	EP 99120289.6				
PRIOR FILING DATE:	1999-10-11				
NUMBER OF SEQ ID NOS:	169				
SOFTWARE:	PatentIn Ver. 2.0				
SEQ ID NO 151					
LENGTH:	440				
TYPE:	PRT				
ORGANISM:	A. fumigatus 32722				

Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 331 DFDHNSMVSIFAL-----GLYNGTEPLSRTSVESAKELDGYASWVVPFGARAYFE 383
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 384 TMOCKSEKPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 430

RESULT 14

US-09-488-265B-6
; Sequence 6, Application US/09488265B
; Patent No. 6720174
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Improved Phytases
; FILE REFERENCE: 5808.200-US
; CURRENT APPLICATION NUMBER: US/09/488,265B
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus 13073
US-09-488-265B-6

Query Match 5.2%; Score 119; DB 4; Length 440;

Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVRAPT-----KATQLMDQVTPDA--- 58
Db 27 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKKYKLVTAIQANATDF 83
Qy 59 -----WPTWPKLGM--LTPRGELIAYLG--HYQORQLVADGLLAKKGCPOSGQVAII 108
Db 84 KGKFAFLKTYNTLTGADDLTPFGGEQLVNSGIKFQRYKALARSVP-----FIR 133
Qy 109 ADVDERTRKTGEAFAAGL-----APDCAITVHTQADTSSPDPLENPLKTGV 154
Db 134 ASGSDRVIASGEKFIQFQOAKLADPGATNRAAPAIISVII-PESETFN-----NTLDHGV 187
Qy 155 COLDNANVTDAIILSRAGSSI-ADFTGHRQTAFRELERVLNFPQSNLCKREKODESCSLT 213
Db 188 C-----TKFEASQLGDEVANFT-----ALFAP--DIRARAEKHLPGVTLT 226
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHOWNTLLSL 271
Db 227 DEDVVSMDMCSFDTVARTSDASQLSPFCQLFTHNE-----W-----KKYNYLQSL 272
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIIMALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 273 --GKYGYGAGNPLGPAQGGIGFTNELIARLTSRSPQDHTSTNSTLVSNPATFPLNATMYV 330
Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 331 DFDHNSMVSIFAL-----GLYNGTEPLSRTSVESAKELDGYASWVVPFGARAYFE 383
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 384 TMOCKSEKPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 430

RESULT 15

US-09-044-718-12
; Sequence 12, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; PASAMONTES, Luis

; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-12

Query Match 5.2%; Score 119; DB 3; Length 449;

Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVRAPT-----KATQLMDQVTPDA--- 58
Db 36 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKKYKLVTAIQANATDF 92
Qy 59 -----WPTWPKLGM--LTPRGELIAYLG--HYQORQLVADGLLAKKGCPOSGQVAII 108
Db 93 KGKFAFLKTYNTLTGADDLTPFGGEQLVNSGIKFQRYKALARSVP-----FIR 142
Qy 109 ADVDERTRKTGEAFAAGL-----APDCAITVHTQADTSSPDPLENPLKTGV 154
Db 143 ASGSDRVIASGEKFIQFQOAKLADPGATNRAAPAIISVII-PESETFN-----NTLDHGV 196
Qy 155 COLDNANVTDAIILSRAGSSI-ADFTGHRQTAFRELERVLNFPQSNLCKREKODESCSLT 213
Db 197 C-----TKFEASQLGDEVANFT-----ALFAP--DIRARAEKHLPGVTLT 235
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHOWNTLLSL 271
Db 236 DEDVVSMDMCSFDTVARTSDASQLSPFCQLFTHNE-----W-----KKYNYLQSL 281
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIIMALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 282 --GKYGYGAGNPLGPAQGGIGFTNELIARLTSRSPQDHTSTNSTLVSNPATFPLNATMYV 339
Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 340 DFDHNSMVSIFAL-----GLYNGTEPLSRTSVESAKELDGYASWVVPFGARAYFE 392
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 393 TMOCKSEKPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 439

Search completed: March 11, 2005, 23:09:24
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 23:04:56 ; Search time 142 Seconds
(without alignments)
1022.054 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACLSRHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844958 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	440	9	US-09-777-566A-2
2	2302	100.0	440	9	US-09-866-379-2
3	2302	100.0	440	14	US-10-034-985-2
4	2302	100.0	440	15	US-10-430-356-2
5	2302	100.0	440	15	US-10-601-319-2
6	2235	97.1	432	9	US-09-866-379-8
7	2235	97.1	432	14	US-10-156-660-4
8	2235	97.1	432	15	US-10-282-122A-43351
9	2235	97.1	432	15	US-10-601-319-8
10	2235	97.1	432	16	US-10-472-317-41
11	2221	96.5	432	15	US-10-284-962-3
12	2221	96.5	432	15	US-10-284-962-14
13	2221	96.5	433	14	US-10-266-041-1
14	2221	96.5	433	14	US-10-266-041-1

14	2197	95.4	432	15	US-10-284-962-5	Sequence 5, Appli
15	2189	95.0	436	14	US-10-156-660-2	Sequence 2, Appli
16	2171	94.3	432	15	US-10-601-319-10	Sequence 10, Appl
17	2165	94.0	432	14	US-10-021-723A-16	Sequence 16, Appl
18	2141	93.0	430	9	US-09-866-379-10	Sequence 10, Appl
19	2070	89.9	412	14	US-10-334-672-1	Sequence 1, Appli
20	2070	89.9	412	14	US-10-334-671-1	Sequence 1, Appli
21	2066	89.7	410	14	US-10-021-723A-13	Sequence 13, Appl
22	2066	89.7	410	14	US-10-021-723A-15	Sequence 15, Appl
23	960.5	41.7	441	14	US-10-021-723A-12	Sequence 12, Appl
24	943.5	41.0	441	15	US-10-282-122A-77792	Sequence 77792, A
25	935.5	40.6	441	14	US-10-021-723A-2	Sequence 2, Appli
26	932.5	40.5	420	14	US-10-021-723A-4	Sequence 4, Appli
27	760	33.0	476	14	US-10-021-723A-10	Sequence 10, Appl
28	611.5	26.6	318	14	US-10-021-723A-14	Sequence 14, Appl
29	476.5	20.7	409	14	US-10-021-723A-8	Sequence 8, Appli
30	472.5	20.5	421	14	US-10-021-723A-6	Sequence 6, Appli
31	195	8.5	426	15	US-10-257-174-44	Sequence 44, Appl
32	195	8.5	426	15	US-10-343-357-6	Sequence 6, Appli
33	121	5.3	465	16	US-10-662-914-10	Sequence 10, Appl
34	119	5.2	439	14	US-10-062-848-3	Sequence 3, Appli
35	119	5.2	439	16	US-10-776-104-3	Sequence 10, Appl
36	119	5.2	440	15	US-10-442-538-106	Sequence 106, App
37	119	5.2	440	15	US-10-442-538-128	Sequence 128, App
38	119	5.2	440	15	US-10-442-538-151	Sequence 151, App
39	119	5.2	449	14	US-10-062-848-12	Sequence 12, Appl
40	119	5.2	449	16	US-10-776-104-12	Sequence 12, Appl
41	119	5.2	465	13	US-10-083-452-8	Sequence 8, Appli
42	119	5.2	465	14	US-10-062-848-78	Sequence 78, Appl
43	119	5.2	465	14	US-10-229-358-6	Sequence 6, Appli
44	119	5.2	465	16	US-10-662-914-4	Sequence 4, Appli
45	119	5.2	465	16	US-10-776-104-78	Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-09-777-566A-2
; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2

Query Match 100.0%; Score 2302; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 28-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
QY 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPKQKQAYGVTLFTSVLFIAGHDTNLNLANLGGALELWNTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPKQKQAYGVTLFTSVLFIAGHDTNLNLANLGGALELWNTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQCMCSLAGFTQ 420
QY 421 INVNEARIPACSLRSHHHHH 440
DB 421 INVNEARIPACSLRSHHHHH 440

RESULT 2
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR FILING DATE: 2000-05-25
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 1999-04-13
; PRIOR FILING DATE: 1999-03-01
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-866-379-2

Query Match 100.0%; Score 2302; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60

DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
QY 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPKQKQAYGVTLFTSVLFIAGHDTNLNLANLGGALELWNTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPKQKQAYGVTLFTSVLFIAGHDTNLNLANLGGALELWNTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQCMCSLAGFTQ 420
QY 421 INVNEARIPACSLRSHHHHH 440
DB 421 INVNEARIPACSLRSHHHHH 440

RESULT 3
US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-034-985-2

Query Match 100.0%; Score 2302; DB 14; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
QY 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKLWLTTPRGELIAYIGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180

```
QY 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPSTVLFIAGHDNTNLNLCGALNLTLPQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPSTVLFIAGHDNTNLNLCGALNLTLPQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWTOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 4
US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2

Query Match 100.0%; Score 2302; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
QY 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTKTGE 120
DB 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
```

```
QY 301 LTPHPQKQAYGVTLPSTVLFIAGHDNTNLNLCGALNLTLPQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPSTVLFIAGHDNTNLNLCGALNLTLPQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWTOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
US-10-601-319-2
; Sequence 2, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-2

Query Match 100.0%; Score 2302; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
QY 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTKTGE 120
DB 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
```

```
QY 301 LTPHPKQAYGVTLPSTVLFIAHDTNLNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPSTVLFIAHDTNLNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 6
US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVERSI370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-866-379-8

Query Match 97.1%; Score 2235; DB 9; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELRLVNPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
QY 301 LTPHPKQAYGVTLPSTVLFIAHDTNLNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
```

```
Db 301 LTPHPKQAYGVTLPSTVLFIAHDTNLNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 7
US-10-156-660-4
; Sequence 4, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-156-660-4

Query Match 97.1%; Score 2235; DB 14; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAIIPLLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELRLVNPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
```

QY 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLANLGALBNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLANLGALBNWTLPGQPDNTPPGGELVPERW 360
QY 361 RRLSDNSQWQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 8
US-10-282-122A-43351
; Sequence 43351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43351
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43351

Query Match 97.1%; Score 2235; DB 15; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
Db 1 MKAILPFLSLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120

Db 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 RQAFRELERVLNFPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNCLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Db 241 EIFLLOQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLANLGALBNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLANLGALBNWTLPGQPDNTPPGGELVPERW 360
QY 361 RRLSDNSQWQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 9
US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A. Robert
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-8

Query Match 97.1%; Score 2235; DB 15; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
Db 1 MKAILPFLSLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDWP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120

```
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
Qy 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 10
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003US1
; CURRENT APPLICATION NUMBER: US/10/472,317
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-472-317-41
```

```
Query Match 97.1%; Score 2235; DB 16; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
```

```
Qy 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 11
US-10-284-962-3
; Sequence 3, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-284-962-3
```

```
Query Match 96.5%; Score 2221; DB 15; Length 432;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQKAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
```


RESULT 12

US-10-284-962-14
; Sequence 14, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-284-962-14

Query Match 96.5%; Score 2221; DB 15; Length 432;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Qy 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQPGQVAIIADVDERTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQPGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAFTGH 180

Qy 181 QTAFARELERVLPFQSNCLNREKQDESLTQALPSELKVSADNVSLSLTAAGVSLASMLT 240
Db 181 QTAFARELERVLPFQSNCLNREKQDESLTQALPSELKVSADNVSLSLTAAGVSLASMLT 240

Qy 241 EIFLLQQAQMPGPGWGRITDSHWNLTLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHWNLTLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIVQSVLFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 13

US-10-266-041-1
; Sequence 1, Application US/10266041
; Publication No. US20030072844A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-10-266-041-1

Query Match 96.5%; Score 2221; DB 14; Length 433;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Qy 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQPGQVAIIADVDERTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQPGQVAIIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAFTGH 180

Qy 181 QTAFARELERVLPFQSNCLNREKQDESLTQALPSELKVSADNVSLSLTAAGVSLASMLT 240
Db 181 QTAFARELERVLPFQSNCLNREKQDESLTQALPSELKVSADNVSLSLTAAGVSLASMLT 240

Qy 241 EIFLLQQAQMPGPGWGRITDSHWNLTLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHWNLTLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIVQSVLFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 14

US-10-284-962-5
; Sequence 5, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-284-962-5

Query Match 95.4%; Score 2197; DB 15; Length 432;
Best Local Similarity 98.1%; Pred. No. 3.3e-184;

Matches 424; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Qy 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTRKTGE 120
Db 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 QTAFARELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOQAQMPPEPGWGRIITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLLOQAQMPPEPGWGRIITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Qy 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 15

US-10-156-660-2
; Sequence 2, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156.660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-156-660-2

Query Match 95.0%; Score 2188; DB 14; Length 436;
Best Local Similarity 97.7%; Pred. No. 2.1e-183;
Matches 425; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Qy 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTRKTGE 120
Db 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 QTAFARELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOQAQMPPEPGWGRIITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Db 241 EIFLLOQAQMPPEPGWGRIITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
Qy 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSH 435
Db 421 IVNEARIPACSLRSH 435

Search completed: March 11, 2005, 23:20:01
Job time : 144 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:58:45 ; Search time 41 Seconds
(without alignments)
1032.570 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLRSHHHHH 440
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235	97.1	432	2 B36733	acid phosphatase (
2	2210	96.0	434	2 H90770	phosphoanhydride p
3	2210	96.0	444	2 D85633	hypothetical prote
4	943.5	41.0	441	2 AC0201	acid phosphatase (
5	545	23.7	413	2 F90773	periplasmic glucos
6	545	23.7	413	2 B85636	periplasmic glucos
7	541	23.5	417	2 S25627	glucose-1-phosphat
8	539	23.4	413	2 JV0087	glucose-1-phosphat
9	526	22.8	414	2 E87316	periplasmic phosph
10	524	22.8	413	2 AG0632	glucose-1-phosphat
11	139.5	6.1	423	1 S06167	acid phosphatase (
12	133	5.8	416	2 T16058	hypothetical prote
13	131	5.7	423	2 A33395	acid phosphatase (
14	122.5	5.3	421	2 S14742	acid phosphatase (
15	117	5.1	386	1 JH0610	acid phosphatase (
16	113.5	4.9	344	2 B89130	protein F52E1.8 [1
17	109	4.7	2554	1 TVFF7L	Kinase-related pro
18	108.5	4.7	10223	2 T30225	polyketide synthas
19	107.5	4.7	457	2 A56925	paired box transcr
20	107	4.6	450	2 A54429	paired box transcr
21	106.5	4.6	1048	1 BVRESC	exonuclease (EC 3.
22	105.5	4.6	459	2 S52250	paired box transcr
23	104.5	4.5	2535	2 AC0304	probable hemolysin
24	102	4.4	638	2 A29440	signal recognition
25	101.5	4.4	397	2 C81716	hypothetical prote
26	101.5	4.4	479	1 JN0890	acid phosphatase (
27	101.5	4.4	524	2 H75530	probable acid-CoA
28	101.5	4.4	1047	2 C85535	ATP-dependent dsDN
29	101.5	4.4	1047	2 G90694	ATP-dependent dsDN

ALIGNMENTS

RESULT 1

B36733
acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N;Alternate names: phosphoanhydride phosphohydrolase, periplasmic; phytase 6
C;Species: Escherichia coli
C;Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C;Accession: B36733; S18018; B64839; A26534; S17960; S33278
R;Dassa, J.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 5497-5500, 1990
A;Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals sign
A;Reference number: A36733; MUID:90368616; PMID:2168385
A;Accession: B36733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-432 <DAS>
A;Cross-references: UNIPROT:P07102; GB:M58708; NID:g145283; PIDN:AAA72086.1; PID:g145283
R;Greiner, R.; Jany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A;Title: Characterization of a phytase from Escherichia coli.
A;Reference number: S18018
A;Accession: S18018
A;Molecule type: protein
A;Residues: 23-33 <GRE>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64839
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-432 <BLAT>
A;Cross-references: GB:AE000200; GB:U00096; NID:g2367111; PIDN:AACT4065.1; PID:g1787215;
A;Experimental source: strain K-12, substrain MG1655
R;Touati, E.; Danchin, A.
Biochimie 69, 215-221, 1987
A;Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phosph
A;Reference number: A26534; MUID:87271766; PMID:3038201
A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-50, 'NAGCHPRRMANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>
A;Cross-references: GB:X05471; NID:g40925; PIDN:CAA29031.1; PID:g40927
R;Dassa, J.; Faihi, H.; Marck, C.; Dion, M.; Kleffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A;Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put
A;Reference number: S17958; MUID:9204231; PMID:1658595
A;Accession: S17960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-17 <DA2>
A;Cross-references: GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:g238659
R;Greiner, R.; Konietzny, U.; Jany, K.D.

nofA protein precu
titin, cardiac mus
hydroxymethylbilan
acid phosphatase (

Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from *Escherichia coli*.
A;Reference number: S33278; MUID:93256556; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31, 'A', 33-34 <GR2>
C;Comment: In addition to CAMP-mediated control, this enzyme is induced when bacterial c naerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mon
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 97.1%; Score 2235; DB 2; Length 432;
Best Local Similarity 99.5%; Pred. No. 6.1e-171;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDWP 60
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDWP 60

QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCQSQGOVAILIADVDERTRKTGE 120
DB 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCQSQGOVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180

QY 181 ROTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLQQAQGMPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLQQAQGMPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

QY 301 LTHPPKQKQAYGVTLPTSVLFIAAGDNTNLNGLGALINLWTLPGQPDNTPPGGELVPERW 360
DB 301 LTHPPKQKQAYGVTLPTSVLFIAAGDNTNLNGLGALINLWTLPGQPDNTPPGGELVPERW 360

QY 361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 2
H90770
phosphoanhydride phosphorylase [imported] - *Escherichia coli* (strain O157:H7, substrain
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; GB:BA000007; PIDN:BA34559.1; PID:g13360596; GSPDB:C
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECa1136

Query Match 96.0%; Score 2210; DB 2; Length 434;
Best Local Similarity 98.4%; Pred. No. 6.1e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 58
DB 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 60

QY 59 WPTWPKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCQSQGOVAILIADVDERTRKT 118
DB 61 WPNWPKLGWLTTPRGGLIAYLGHYQRLVADGLLTKGCPQPGQVAILIADVDERTRKT 120

QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 178
DB 121 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 180

QY 179 GHRQTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLAS 238
DB 181 GHRQTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLAS 240

QY 239 LTFEILQQAQGMPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIM 298
DB 241 LTFEILQQAQGMPEPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIM 300

QY 299 AALTTPHPKQKQAYGVTLPTSVLFIAAGDNTNLNGLGALINLWTLPGQPDNTPPGGELVPE 358
DB 301 IALTTPHPKQKQAYGVTLPTSVLFIAAGDNTNLNGLGALINLWTLPGQPDNTPPGGELVPE 360

QY 359 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 418
DB 361 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 420

QY 419 TQIVNEARIPACSL 432
DB 421 TQIVNEARIPACSL 434

RESULT 3
D85633
hypothetical protein appA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93:
C;Species: *Escherichia coli*
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; GB:AE005174; NID:g12514245; PIDN:AAG55528.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Query Match 96.0%; Score 2210; DB 2; Length 444;
Best Local Similarity 98.4%; Pred. No. 6.3e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 58
DB 11 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 70

QY 59 WPTWPKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCQSQGOVAILIADVDERTRKT 118
DB 71 WPNWPKLGWLTTPRGGLIAYLGHYQRLVADGLLTKGCPQPGQVAILIADVDERTRKT 130

QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 178
DB 131 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 190

Db 366 ELWQPNDHQRVYAVKMFYQWMDQLRNAEKLDLKNNPAGIISVAVAGCENNGDDKCELD 425

Qy 417 GFTQIVNEARIPACSL 432
| : : : | : | :
Db 426 TFQKKVAKVIEPACHI 441

RESULT 5
F90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90773
F:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ihii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90773
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <HAY>
A:Cross-references: UNIPROT:Q8XBZ6; GB:BA000007; PIDN:BAH34581.1; PID:g13360618; GSPDB:G1
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECa1158

Query Match 23.7%; Score 545; DB 2; Length 413;
Best Local Similarity 32.1%; Pred No. 8.9e-36;
Matches 135; Conservative 76; Mismatches 180; Indels 30; Gaps 10;

Qy 13 IPLTPQSAFAOSEPE-LKLESVIVSRHGVRAP-TKATOLMODVTPDAMPTWPKLGLWT 70
Db 13 VVLLASNAQAQTVPEGYQLQVLMMSRHLRAPLANNGSVLEQSTENKWPENDVPGQLT 72

Qy 71 PRGELIAYLGHYQORLVDAGLLAKKCPQSGQVAIIADVDERTKRTGEAFAGLAPDC 130
Db 73 TKGGVLEYIMGHYKREWLAEGMVKSGECPPPTTVYVANSQRTVATQAQFFITGAPPG 132

Qy 131 AITVHTQADTSSPPFLNPLKTVGCOLDNANVTDAILSRAGGSIAADFTGHRQTFARELER 190
Db 133 DIPVHQEKMGTMDDTFNPVTD----DJAASEQAAVAMEKLSKL--QLTDSYQLLEK 186

Qy 191 VLNFPQSNLCILKREKQDESCSLTQALPSSELKVSADNVSLTGAVSLASMLTEIFLLQQAQG 250
Db 187 IVNYKDSFAC---KEKQCCSLVDGKNTFSAKYQEPGVSGPLKVGNSLVDAFTLQYEG 242

Qy 251 MP--EPGWGRITDHSQWNTLLSHNAQYVLLQRTTEVARSRAATPLLDLIMAAALTTPHPQK 308
Db 243 FPMQCVAMGEIKSDQMWKLVSLKNGYQDSLFTSPSEARNVAKPLVSYIDKAL----- 295

Qy 309 QAYGVTLPTS---VLFIAGHDNLNANLGALELN-WTLPGQPDNTPPGELVFERWRLS 364
Db 296 ----VTDRTPAKTVLVGHDSNIAASLTALDFKPYQLHDQDERTPIGKIVFQWRDVK 351

Qy 365 DNSQWIVSLVFQTLQMRDKTPLSLNTPPGSEVKLTLAGCBERNAGCMCSLAGFTQVNE 424
Db 352 ANRDLMKIEYYQSAEQLRNADALTLPAPQVRLLELSC-PIDADGFCPMDKFDVSLNE 410

Qy 425 A 425
|
Db 411 A 411

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85636
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001

QY 2 KAILIPF-----LSLPIPTPQSAFAQSBEFLKESVIVSRHGVRAPTATQMDDVTP 56
:
:
:
Db 8 QAALLQLFGMLCTVMPPIQARS-----LRFTLLRYHGRDRSPVKT-----YPKDP 53
:
:
:
QY 57 DAWPTWPKLGWLTPGGELIAVLGHVQRRLVADGLLAKKGCPQSGOVAIADVD-ERT 115
:
:
:
Db 54 YQBERWPQGFGQTKEGMIQHWEIGALQRY--HGFL---NTSYHRQEVYVSRSTDPRD 108
:
:
:
QY 116 RKTGEAFAAGLAPDCAITVHTQADTSPPDFNPLKTVGCQLDNANVTDAILSRAGGSIA 175
:
:
:
Db 109 LMSAEANLAGLPP-----NEVQHFSPNISMQPI-----PVHTVPITE----- 146
:
:
:
QY 176 DFTGHRGTAFRELERNLVLPFQSNLCRKSKQDSCLTQAALPELKVSADVSLTGAVSL 235
:
:
:
Db 147 -----DRLLKFPGLGPCRYPEQLQNE----TROTPEQNRISIQNAOFLNMVAN 189
:
:
:
QY 236 ASMLTEIFL-----LOOAQMPERPCHGRITSDSHQWNTHLLSLHNAQFYLLQRTPE 284
:
:
:
Db 190 ETGLTNVTLETINWNYDTLFCETGTHLLLPFWASPTQVQLSKDKFSFLFLGIHEQVQ 249
:
:
:
QY 285 VARSRAITLLIDTMAALTTPHPPOKAQYGVTLPSTSVLFIAGHDTNLNMGCALELNMTLPG 344
:
:
:
Db 250 KARLOGGVLLAQILKNLTLMATTSSQP-----PKLLVYSADHTTLVALQNALNVY----- 298
:
:
:
QY 345 QPDNTPGGELVFERRRRISDNSQTVQSLVFOOTLOOMRDKTPLSLNTNPGEVKLTLACG 404
:
:
:
Db 299 NGKOAPVASYACHIPELYQ--EDNGNF-SVEMYFRNSDK--KAPWPLILP-----GC 343
:
:
:
QY 405 EERNAOGMCSLAGTOIVNEARIP 428
:
:
:
Db 344 PHR-----CPLQDFLEL-TPEVIP 361
:
:
:

RESULT 15

JH0610
acid phosphatase (EC 3.1.3.2) ACP precursor [validated] - human
N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphatase; Homo sapiens (man)
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence revision 01-Dec-1995 #text change 09-Jul-2004
C;Accession: JH0610; J50693; A38608; S01331; A32419; S11147; S38863; S41251; S1
R;Sharief, F.S.; Li, S.S.L.
A;Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992
A;Title: Structure of human prostatic acid phosphatase gene.
A;Reference number: JH0610; MUID:92272747; PMID:1375464
A;Accession: JH0610
A;Molecule type: DNA
A;Residues: 1-386 <SHA>
A;Cross-references: UNIPROT:P15309; GB:M97580; GB:M97581; GB:M97582; GB:M97583;
A;Accession: J50693
A;Molecule type: mRNA
A;Residues: 1-386 <SH3>
A;Cross-references: GB:M97589; NID:g189611; PIDN:AAA60021.1; PID:g189613
R;Van Etten, R.L.; Davidsen, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.
J. Biol. Chem. 266, 2313-2319, 1991
A;Title: Covalent structure, disulfide bonding, and identification of reactive
A;Reference number: A38608; MUID:91115848; PMID:1989985
A;Accession: A38608
A;Molecule type: mRNA
A;Residues: 1-386 <VAN>
A;Cross-references: GB:M34840; NID:g189620; PIDN:AAA69694.1; PID:g189621
A;Note: part of this sequence, including the amino end of the mature protein, w
R;Viikko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Huhtala, M.L.
FEBS Lett. 236, 275-281, 1988
A;Title: Molecular cloning and sequence analysis of cDNA encoding human prostat
A;Reference number: S01331; MUID:88312981; PMID:2842184
A;Accession: S01331
A;Molecule type: mRNA
A;Residues: 1-14, 'A', '16', 'ASC', '20', 'CF', '23', 'C', '25-65', 'WIWPTHPA', '74-211', 'A', '213-38'
A;Cross-references: EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID:g28322
A;Note: part of this sequence, including the amino end of the mature protein, w
R;Sharief, F.S.; Lee, H.; Leudermaun, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.;
Biochem. Biophys. Res. Commun. 160, 79-86, 1989

[illegible]

Search completed: March 11, 2005, 23:08:35
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:58:05 ; Search time 172 Seconds
(without alignments)
1309.970 Million cell updates/sec

Title: US-10-034-985-2

Perfect score: 2302

Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2235	97.1	432	1	PPA_ECOLI	P07102 escherichia
2	2231	96.9	432	2	Q8GN88	Q8gn88 escherichia
3	2229	96.8	432	2	Q8RKD6	Q8rkd6 escherichia
4	2229	96.8	432	2	Q8RKD7	Q8rkd7 escherichia
5	2229	96.8	432	2	Q8RKD8	Q8rkd8 escherichia
6	2229	96.8	432	2	Q8RKD9	Q8rkd9 escherichia
7	2228	96.8	432	2	Q8RKD5	Q8rkd5 escherichia
8	2225	96.7	432	2	Q83RW2	Q83rw2 shigella fl
9	2222	96.5	442	2	Q8RK08	Q8rk08 escherichia
10	2221	96.5	432	2	Q7UD08	Q7ud08 shigella fl
11	2218	96.4	432	2	Q8C75	Q8cw75 escherichia
12	2213	96.1	446	2	Q8C75	Q8cw75 escherichia
13	2210	96.0	434	2	Q7AFW3	Q7afw3 escherichia
14	2210	96.0	444	2	Q8XC29	Q8xc29 escherichia
15	1374.5	59.7	433	2	Q676V7	Q676v7 citrobacter
16	1077	46.8	444	2	Q676V7	Q6u677 obesumbacte
17	1054.5	45.8	444	2	Q67AQ8	Q6ta08 obesumbacte
18	943.5	41.0	441	2	Q66R3	Q66r3 yersinia ps
19	943.5	41.0	441	2	Q82FP6	Q82fp6 yersinia ps
20	545	23.7	413	2	Q8XB26	Q8xb26 escherichia
21	541	23.5	417	1	AGP_PRORE	P19926 providencia
22	539	23.4	413	1	AGP_ECOLI	Q52309 escherichia
23	538	23.4	413	2	Q7UD02	Q7ud02 shigella fl
24	538	23.4	421	2	Q83RV6	Q83rv6 shigella fl
25	529	23.0	413	1	AGP_SALTY	O33921 salmonella
26	526	22.8	414	2	Q9AAQ4	Q9aaq4 caulobacter
27	524	22.8	413	2	Q8Z7P1	Q8z7p1 salmonella
28	523	22.7	413	2	Q6EV19	Q6ev19 enterobacte
29	505	21.9	392	2	Q8P76	Q8pp76 xanthomonas
30	505	21.9	443	2	Q8P330	Q8p330 xanthomonas
31	480.5	20.9	435	2	Q8PP53	Q8pf53 xanthomonas

32	467.5	20.3	435	2	Q6CZP4	O6czf4 erwinia car
33	433.5	18.8	428	2	Q8GD20	Q8gd20 pseudomonas
34	392	17.0	318	2	Q8CW72	Q8cw72 escherichia
35	373	16.2	414	2	Q8PPX6	Q8ppx6 xanthomonas
36	359	15.6	421	2	Q84CN9	Q84cn9 klebsiella
37	353	15.3	421	2	Q7WSY1	Q7wsy1 klebsiella
38	221.5	9.6	419	2	Q8VQS2	Q8vq2 klebsiella
39	195	8.5	426	2	Q9BZG2	Q9bzg2 homo sapien
40	150	6.5	531	2	Q9GNZ3	Q9gnz3 leishmania
41	140	6.1	423	1	PPAL_HUMAN	P1117 homo sapien
42	140	6.1	542	2	Q9ULI2	Q9ula2 leishmania
43	139.5	6.1	333	2	Q9BZG3	Q9bzg3 homo sapien
44	139	6.0	462	2	Q7QC05	Q7qc05 anopheles g
45	137.5	6.0	406	2	Q66L09	Q66l09 xenopus lae

ALIGNMENTS

RESULT 1

ID	PPA_ECOLI	STANDARD;	PRT;	432 AA.
AC	P07102;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Periplasmic appa protein precursor [Includes: Phosphoanhydride			
DE	phosphohydrolase [EC 3.1.3.2] (pH 2.5 acid phosphatase) (AP); 4-			
DE	phytase [EC 3.1.3.26]]			
GN	Names-appa; OrderedLocusNames=b0980;			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 23-29.			
RC	STRAIN=K12;			
RX	MEDLINE=90368616; PubMed=2168385;			
RA	Dassa J., Marck C., Boquet P.L.;			
RT	"The complete nucleotide sequence of the Escherichia coli gene appa			
RT	reveals significant homology between the pH 2.5 acid phosphatase and			
RT	glucose-1-phosphatase."			
RL	J. Bacteriol. 172:5497-5500 (1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.;"			
RL	Science 277:1453-1474 (1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RX	MEDLINE=970611202; PubMed=8905232;			
RA	Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,			
RA	Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,			
RA	Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,			
RA	Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,			
RA	Sampe G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,			
RA	Yano M., Horiuchi T.;			
RT	"A 718-kb DNA sequence of the Escherichia coli K-12 genome			
RT	corresponding to the 12.7-28.0 min region on the linkage map.;"			
RL	DNA Res. 3:137-155 (1996).			
RN	[4]			
RP	SEQUENCE OF 1-112 FROM N.A.			
RX	MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;			
RA	Touati E., Danchin A.;			
RT	"The structure of the promoter and amino terminal region of the pH 2.5			
RT	acid phosphatase structural gene (appa) of E. coli: a negative control			
RT	of transcription mediated by cyclic AMP."			

RL Biochimie 69:215-221(1987).
RN [5]
RP SEQUENCE OF 1-17 FROM N.A.
RC STRAIN=K12;
RA MEDLINE=92049231; PubMed=1658595;
RX Dassa J., Fsihi H., Marck C., Dion M., Kieffer-Bontemps M.,
RT Boquet P.L.;
RT "A new oxygen-regulated operon in Escherichia coli comprises the genes
RT for a putative third cytochrome oxidase and for pH 2.5 acid
RT phosphatase (appa).";
RL Mol. Gen. Genet. 229:341-352(1991).
RN [6]
RP CHARACTERIZATION, AND SEQUENCE OF 23-34.
RA Greiner R., Jany K.D.;
RT "Characterization of a phytase from Escherichia coli.";
RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN [7]
RP CHARACTERIZATION, AND SEQUENCE OF 23-35.
RX MEDLINE=93256556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA Greiner R., Konietzny U., Jany K.D.;
RT "Purification and characterization of two phytases from Escherichia
RT coli.";
RL Arch. Biochem. Biophys. 303:107-113(1993).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=20161462; PubMed=10696472;
RA Golovan S., Wang G., Zhang J., Forsberg C.W.;
RT "Characterization and overproduction of the Escherichia coli appA
RT encoded bifunctional enzyme that exhibits both phytase and acid
RT phosphatase activities.";
RL Can. J. Microbiol. 46:59-71(2000).
RN [9]
RP MUTAGENESIS.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RT van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA Lim D., Golovan S., Forsberg C.W., Jia Z.;
RT "Crystal structures of Escherichia coli phytase and its complex with
RT phytate.";
RL Nat. Struct. Biol. 7:108-113(2000).
CC -1- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
CC phosphate.
CC -1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
CC myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: In addition to CAMP-mediated control, this enzyme is
CC induced when bacterial cultures reach stationary phase; its
CC synthesis is triggered by phosphate starvation or a shift from
CC aerobic to anaerobic conditions.
CC -1- SIMILARITY: Belongs to the histidine acid phosphatase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M58708; AAC72086.1; -;
CC EMBL; U00096; AAC74065.1; -;
CC EMBL; D90735; BAA35745.1; -;
CC EMBL; X05471; CAA29031.1; -;
CC EMBL; M58708; -; NOT_ANNOTATED_CDS.
CC EMBL; S63811; AAB20286.1; -;
CC PIR; B36733; B36733.
DR PDB; 1DKL; X-ray; A/B=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.
DR PDB; 1DKP; X-ray; A=23-432.
DR PDB; 1DKQ; X-ray; A=23-432.
DR PDB; 1DKR; X-ray; A=23-432.
DR PDB; 1DKS; X-ray; A=23-432.
DR PDB; 1DKT; X-ray; A=23-432.
DR PDB; 1DKU; X-ray; A=23-432.
DR PDB; 1DKV; X-ray; A=23-432.
DR PDB; 1DKW; X-ray; A=23-432.
DR PDB; 1DKX; X-ray; A=23-432.
DR PDB; 1DKY; X-ray; A=23-432.
DR PDB; 1DKZ; X-ray; A=23-432.
DR PDB; 1DKA; X-ray; A=23-432.
DR PDB; 1DKB; X-ray; A=23-432.
DR PDB; 1DKC; X-ray; A=23-432.
DR PDB; 1DKD; X-ray; A=23-432.
DR PDB; 1DKE; X-ray; A=23-432.
DR PDB; 1DKF; X-ray; A=23-432.
DR PDB; 1DKG; X-ray; A=23-432.
DR PDB; 1DKH; X-ray; A=23-432.
DR PDB; 1DKI; X-ray; A=23-432.
DR PDB; 1DKJ; X-ray; A=23-432.
DR PDB; 1DKK; X-ray; A=23-432.
DR PDB; 1DKL; X-ray; A=23-432.
DR PDB; 1DKM; X-ray; A=23-432.
DR PDB; 1DKN; X-ray; A=23-432.
DR PDB; 1DKO; X-ray; A=23-432.

```

FT STRAND 354 362
FT TURN 363 366

Query Match
Best Local Similarity 97.1%; Score 2235; DB 1; Length 432;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAIIIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKLGWLTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120
DB 61 TWPVKLGWLTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIILSRAGGSIADFTGH 180

QY 181 ROTAFRELERVLNPPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNPPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
DB 241 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWISQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWISQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420

QY 421 INVNEARIPACSL 432
DB 421 INVNEARIPACSL 432

RESULT 3
Q8KRD6 PRELIMINARY; PRT; 432 AA.
AC Q8KRD6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., March C., Boquet P.-L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR HSSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; AF86C41EA6193AC5 CRC64;

RESULT 2
Q8GN88 PRELIMINARY; PRT; 432 AA.
AC Q8GN88;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE AppA.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chen Y., Zhu Z., Zhang Z., He J.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF537219; AAN28334.1; -.
DR HSSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 432 AA; 47055 MW; 5B355D76E7377737 CRC64;

Query Match
Best Local Similarity 96.9%; Score 2231; DB 2; Length 432;
Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Query Match      96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.8e-164;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAIIPLFSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAIIPLFSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELERVLNFPQSNLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQGMPEPGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOAQGMPEPGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 4
Q8RKD7 Q8RKD7 PRELIMINARY; PRT; 432 AA.
AC Q8RKD7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.I.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL; L03373; AAA00005.1; --
DR HSP; P07102; IDKM
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR00560; HisAc phosphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.

Query Match      96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.8e-164;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAIIPLFSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAIIPLFSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELERVLNFPQSNLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQGMPEPGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOAQGMPEPGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWITQVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 5
Q8RKD8 Q8RKD8 PRELIMINARY; PRT; 432 AA.
AC Q8RKD8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.I.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
DR EMBL; L03372; AAA00004.1; --
DR HSP; P07102; IDKM.

```

```

DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc phosphatase.
DR Pfam; PF00328; Acid phosphatase A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; 9885536857FCCFB5 CRC64;

Query Match          96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.8e-164;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLNFPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSLNTPPEGVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSLNTPPEGVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 6
Q8RKE0 PRELIMINARY; PRT; 432 AA.
AC Q8RKE0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."

J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03370; AAA00002.1; -.
DR HSP; P07102; IDRM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc phosphatase.
DR Pfam; PF00328; Acid phosphatase A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; 5BB0632D4682EABF CRC64;

Query Match          96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.8e-164;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLNFPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSLNTPPEGVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSLNTPPEGVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 7
Q8RKD9 PRELIMINARY; PRT; 432 AA.
AC Q8RKD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Zhou M.M.,

```

RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 Escherichia coli acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836(1992).
 DR EMBL; L03371; AAA00003.1; -.
 DR HSP; P07102; IDKQ.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; Hisac_phspase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 47033 MW; 9F29B9DF9C368175 CRC64;
 Query Match 96.88; Score 2228; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 2.2e-164;
 Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 QY 181 RQAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 RQAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTHPPKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWLPQPDNTPPGGELVFERW 360
 DB 301 LTHPPKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWLPQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
 DB 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 8
 ID Q8RKDS PRELIMINARY; PRT; 432 AA.
 AC Q8RKDS;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Periplasmic phosphoanhydride phosphohydrolase.
 GN Name=appA;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Boquet P.-L.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase and
 glucose-1-phosphatase.";
 RL J. Bacteriol. 172:5497-5500(1990).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=93054596; PubMed=1429631;
 RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
 RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 Escherichia coli acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836(1992).
 DR EMBL; L03375; AAA00007.1; -.
 DR HSP; P07102; IDKM.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; Hisac_phspase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 46990 MW; 951F393EA9A1A47C CRC64;
 Query Match 96.7%; Score 2225; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 3.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 QY 181 RQAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 RQAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTHPPKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWLPQPDNTPPGGELVFERW 360
 DB 301 LTHPPKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWLPQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
 DB 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 9
 ID Q83RW2 PRELIMINARY; PRT; 442 AA.
 AC Q83RW2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name=appA; OrderedLocusNames=SF0982;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL: AE015127; AAN42610.1; -.
DR HSSP: P07102; 1DK.
DR GO: GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro: IPR000560; HisAc_phspatse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
KW Complete proteome.
SQ SEQUENCE 442 AA; 48183 MW; 20BCE2C454254680 CRC64;

Query Match 96.5%; Score 2222; DB 2; Length 442;
Best Local Similarity 98.8%; Pred. No. 6.6e-164;
Matches 427; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 11 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 70
Qy 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Db 71 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 130
Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 131 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 190
Qy 181 RQAFRELERVLNFPQSNCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 191 RQVRELERVLNFPQSNCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 250
Qy 241 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 251 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 310
Qy 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWLTPGQPDNTPPGGELVFERW 360
Db 311 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWLTPGQPDNTPPGGELVFERW 370
Qy 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMCSLAGFTQ 420
Db 371 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMCSLAGFTQ 430
Qy 421 INVNEARIPACSL 432
Db 431 INVNEARIPACSL 442

RESULT 10
Q6RK08
ID Q6RK08 PRELIMINARY; PRT; 432 AA.
AC Q6RK08
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acid phosphatase/phytase 2.
GN Name: appA2;
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99194564; PubMed=10092520; DOI=10.1006/bbrc.1999.0361;
RA Rodriguez E., Han Y., Lei X.G.;
RT "Cloning, sequencing, and expression of an *Escherichia coli* acid
RT phosphatase/phytase gene (appA2) isolated from pig colon.";

RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Lei X.G.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY496073; AAR87658.1; -.
DR GO: GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro: IPR000560; HisAc_phspatse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
FT CHAIN 20 432
SQ SEQUENCE 432 AA; 47041 MW; 71B0E5ED2EA2674 CRC64;

Query Match 96.5%; Score 2221; DB 2; Length 432;
Best Local Similarity 99.1%; Pred. No. 7.7e-164;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Db 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 RQAFRELERVLNFPQSNCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQVRELERVLNFPQSNCLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWLTPGQPDNTPPGGELVFERW 360
Db 301 LTHPPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWLTPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMCSLAGFTQ 420
Qy 421 INVNEARIPACSL 432
Db 421 INVNEARIPACSL 432

RESULT 11
Q7UD08
ID Q7UD08 PRELIMINARY; PRT; 432 AA.
AC Q7UD08
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
GN Name: appA; OrderedLocustNames=S1048;
OS *Shigella flexneri*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;

"Complete genome sequence and comparative genomics of Shigella flexneri serotype 2a strain 2457T";
Infect. Immun. 71:2775-2786 (2003).
EMBL; AE016981; AAF16495.1; -.
HSP; P07102; IDK.
GO; GO:0003993; F:acid phosphatase activity; IEA.
InterPro; IPR000560; HisAc_pshptse.
PFam; PF00328; Acid_phosphatase_1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SEQUENCE 432 AA; 47136 MW; 200442239BC66DFC CRC64;
Query Match 96.4%; Score 2218; DB 2; Length 432;
Best Local Similarity 98.6%; Pred. No. 1.3e-163;
Matches 426; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTATQMQDVTDPDWP 60
DB 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTATQMQDVTDPDWP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSGQVAVIADVDERTKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSGQVAVIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELRVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELRVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQMPGPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSATPLLDLIMAA 300
DB 241 EIFLLOAQMPGPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSATPLLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWTLPGQDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWTLPGQDNTPPGGELVFERW 360
QY 361 RRLSDNSQWTVSLVFTQMQRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWTVSLVFTQMQRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 INNEARIPACSL 432
DB 421 INNEARIPACSL 432
RESULT 12
Q8CW75 PRELIMINARY; PRT; 446 AA.
AC Q8CW75 PRELIMINARY; PRT; 446 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic appA protein.
GN Name=appA; OrderedLocusNames=c1121;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=2238234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rossch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016758; AAN79589.1; -.
DR HSP; P07102; IDK.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_pshptse.
DR PFam; PF00328; Acid_phosphatase_1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 446 AA; 48605 MW; F1308CD691DB9F49 CRC64;
Query Match 96.1%; Score 2213; DB 2; Length 446;
Best Local Similarity 98.4%; Pred. No. 3.4e-163;
Matches 425; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTATQMQDVTDPDWP 60
DB 11 MKAILPFLSLLIPLTPKSAFAQSEPELKLESVIVSRHGVRAPTATQMQDVTDPDWP 70
QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSGQVAVIADVDERTKTGE 120
DB 71 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLTKKGCPCQSGQVAVIADVDERTKTGE 130
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
DB 131 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 190
QY 181 ROTAFRELRVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 191 ROTAFRELRVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 250
QY 241 EIFLLOAQMPGPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSATPLLDLIMAA 300
DB 251 EIFLLOAQMPGPGWGRITDTSQWNTLLSHNAQFYLLQRTPEVARSATPLLDLIMAA 310
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWTLPGQDNTPPGGELVFERW 360
DB 311 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSNLWTLPGQDNTPPGGELVFERW 370
QY 361 RRLSDNSQWTVSLVFTQMQRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 371 RRLSDNSQWTVSLVFTQMQRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 430
QY 421 INNEARIPACSL 432
DB 431 INNEARIPACSL 442
RESULT 13
Q7AFW3 PRELIMINARY; PRT; 434 AA.
AC Q7AFW3 PRELIMINARY; PRT; 434 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Phosphoanhydride phosphorylase.
GN OrderedLocusNames=ECs1136;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Havaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).
DR EMBL; AP002554; BAB34559.1; -.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.

```
DR InterPro: IPR000560; HisAc_phaphtse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 434 AA; 47337 MW; F197DF7D1869F9C4 CRC64;

Query Match 96.0%; Score 2210; DB 2; Length 434;
Best Local Similarity 98.4%; Pred. No. 5.5e-163;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 MKAILLPFLSLPLTPQSAFAQS--EPKLKESVIVSRHGVRAPTKATQLMQDVTTPDA 58
DB 1 MKAILLPFLSLPLTPQSAFAQSEPEPELKESVIVSRHGVRAPTKATQLMQDVTTPDA 60
QY 59 WPTWPKVGLWLTPRGGELIAYLGHYQRLVADGLLAKKGCPOSQGVAIADYDERTKT 118
DB 61 WPNWPKVGLWLTPRGGELIAYLGHYQRLVADGLLTKGCPQPGQVAIADYDERTKT 120
QY 119 GEFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFT 178
DB 121 GEFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFT 180
QY 179 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLLTGAVSLASM 238
DB 181 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLLTGAVSLASM 240
QY 239 LTFEFLFLLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
DB 241 LTFEFLFLLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 300
QY 299 AALTTPHPKQKQAVGVTLPSTSVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 358
DB 301 IALTTPHPKQKQAVGVTLPSTSVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 360
QY 359 RWRLSDNSQWISVLFQTLQOMRDKTLPISLNTTPGCVKLTLAGCEERNAQMCSLAGF 418
DB 361 RWRLSDNSQWISVLFQTLQOMRDKTLPISLNTTPGCVKLTLAGCEERNAQMCSLAGF 420
QY 419 TQIVNEARIPACSL 432
DB 421 TQIVNEARIPACSL 434

RESULT 14
Q8XC29 PRELIMINARY; PRT; 444 AA.
AC Q8XC29;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
DE periplasmic.
GN Name=appa; OrderedLocusNames=z1397;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., Ivi, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Goettack E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouais K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR ENBL; AE005292; AAG55528.1; -.
DR PIR; D85633; D85633.
DR PIR; H90770; H90770.
DR HSSP; P07102; 1DKL.

CO: CO:0003993; P: acid phosphatase activity; IEA.
Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 444 AA; 48388 MW; A9AA7E6653AF247E CRC64;

Query Match 96.0%; Score 2210; DB 2; Length 444;
Best Local Similarity 98.4%; Pred. No. 5.7e-163;
Matches 427; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 1 MKAILLPFLSLPLTPQSAFAQS--EPKLKESVIVSRHGVRAPTKATQLMQDVTTPDA 58
DB 11 MKAILLPFLSLPLTPQSAFAQSEPEPELKESVIVSRHGVRAPTKATQLMQDVTTPDA 70
QY 59 WPTWPKVGLWLTPRGGELIAYLGHYQRLVADGLLAKKGCPOSQGVAIADYDERTKT 118
DB 71 WPNWPKVGLWLTPRGGELIAYLGHYQRLVADGLLTKGCPQPGQVAIADYDERTKT 130
QY 119 GEFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFT 178
DB 131 GEFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFT 190
QY 179 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLLTGAVSLASM 238
DB 191 GHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLLTGAVSLASM 250
QY 239 LTFEFLFLLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
DB 251 LTFEFLFLLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 310
QY 299 AALTTPHPKQKQAVGVTLPSTSVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 358
DB 311 IALTTPHPKQKQAVGVTLPSTSVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 370
QY 359 RWRLSDNSQWISVLFQTLQOMRDKTLPISLNTTPGCVKLTLAGCEERNAQMCSLAGF 418
DB 371 RWRLSDNSQWISVLFQTLQOMRDKTLPISLNTTPGCVKLTLAGCEERNAQMCSLAGF 430
QY 419 TQIVNEARIPACSL 432
DB 431 TQIVNEARIPACSL 444

RESULT 15
Q676V7 PRELIMINARY; PRT; 433 AA.
AC Q676V7;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Phytase.
GN Name=phyA;
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VKPM B-4090;
RA Zinin N.V., Syneokii S.P.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AY390262; AAR89622.1; -.
DR InterPro: IPR000560; HisAc_phaphtse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 433 AA; 48506 MW; A7923288F2FCFD44 CRC64;

Query Match 59.7%; Score 1374.5; DB 2; Length 433;
Best Local Similarity 61.3%; Pred. No. 3.8e-98;
Matches 265; Conservative 51; Mismatches 115; Indels 1; Gaps 1;
```

